BEST AVAILABLE COPY

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(71) Applicant: BIOS GROUP LP [US/US]; 317 Paseo de Peralta, Santa Fe, NM 87501 (US).

(72) Inventor: KAUFFMAN, Stuart, A.; 1811 S. Camino Cruz Blanco, Santa Fe, NM 87505 (US).

(74) Agents: MORRIS, Francis, E. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).

[Continued on next page]

(54) Title: A METHOD AND SYSTEM TO SYNTHESIZE PORTFOLIOS OF GOODS, SERVICES OR FINANCIAL INSTRUMENTS
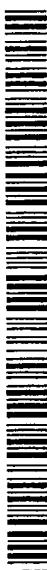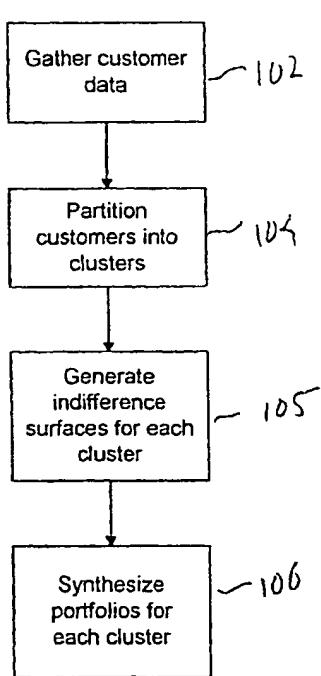
(57) Abstract: The present invention includes methods and systems for dynamically synthesizing custom portfolios of goods, services or financial instruments for clusters of customers from preference data is gathered (102), next, customers are clustered into clusters of similar customers (104), subsequently indifference or utility surfaces are determined that represent the landscape of customer preferences(105), and finally, custom and optimum portfolios are synthesized from the indifference surface and, preferably, historical data concerning the goods, services or financial instruments (106). The present invention also includes computer systems, preferably network-based, distributed systems, that implement the methods of the invention.

100

Gather customer data — 102

Partition customers into clusters — 104

Generate indifference surfaces for each cluster — 105

Synthesize portfolios for each cluster — 106

WO 01/03046 A1

A METHOD AND SYSTEM TO SYNTHESIZE
PORTFOLIOS OF GOODS, SERVICES OR
FINANCIAL INSTRUMENTS

5

FIELD OF THE INVENTION

The present invention relates generally to the synthesis of custom portfolios of goods, services or financial instruments for clusters of customers determined to have similar preferences, in particular to the synthesis of custom portfolios of insurance services for

10   clusters of customers each of whom has insufficient assets for individually customized insurance services.

DESCRIPTION OF RELATED ART

Investment analysis firms, brokerage firms and investment bankers typically provide

15   custom portfolio management to wealthy customers. Specifically, these firms obtain investment information from their wealthy customers, including for example, target return, tolerable risk, time horizon, preferred allocation, tax considerations, and so forth. From this information, these firms synthesize a custom portfolio of stocks, bonds, financial instruments, etc. Because of the expense associated with the custom portfolio management,

20   these firms typically do not offer this service to their other less-wealthy customers.

Insurance companies also do not offer custom insurance programs to their typical customers. For instance, a customer cannot typically acquire insurance on some household goods, such as computer equipment and/or expensive jewelry, while leaving uninsured other goods of less importance or value. Instead, each customer must choose from a fixed

25   and limited number of programs, even if each of the offered programs results in insurance services wasteful to the customer, because, for example, they require insurance of goods for which insurance is not sought in order to insure those goods for which insurance is desired.

More generally, there are numerous other economic or market contexts known

30   where customization of goods and services that are routinely available to wealthier customers or businesses is simply not available to average customers or businesses. The expense of such customization exceeds the likely rewards obtainable from any average customer or business. This results in sub-optimal utility or satisfaction for each individual customer. Providing increased utility to each overage customer or business has not

35   heretofore been exploited because it has been thought that likely expenses outweigh possible returns.

Accordingly, there exists a need for methods and systems that dynamically synthesize custom portfolios of goods, services or financial instruments for average customers or businesses, so that individual customers will obtain greater utility and value than possible with standardized offerings heretofore available in the marketplace.

5

SUMMARY OF THE INVENTION

Therefore, the objects of the present invention are to remedy these defects in the prior art by providing such customized offerings of goods, services, or financial instruments to individual customers or businesses of all purchasing power or size, offerings that

10 necessarily have greater utility than limited standardized offerings available heretofore. These objects are achieved by methods and systems based on novel and original uses of preference data obtained from each individual customer to automatically synthesize such customized portfolios. Individual elements of a portfolio are typically provided by one of more suppliers, for example by manufacturers of goods, providers of insurance services, or

15 brokers or issuers of financial instruments. Complete portfolios can be provided by the primary offerors of the portfolio elements, or by brokers of or dealers in the portfolio elements, or by other market arrangements. Importantly, these automatic systems and methods are effective and are of low cost, allowing the profitable provision of advantageous, customized portfolios widely in the marketplace. The present invention

20 thereby makes possible new and innovative services in the marketplace.

Methods of the present invention start by gathering customer or business preference data. Generally, this data reflects the preferences, or the values, or the utilities of certain goods, services or financial instruments selected from a universe of goods, services, or instruments and for a set of potential customers of businesses. For example, in the case of

25 insurance services, the preference data can represent particular items some customer wishes insured, their economic values, their personal values, and so forth. In the case of financial instruments, the preference data can represent customer wishes concerning the type of instrument, its past risk and reward, expectations for future risk and reward, the geographic area or economic field from which the instrument derives value, and so forth. In the case of

30 goods, especially complex goods, the preference data can represent customer wishes for various combinations of features available with the goods. For example, for automobiles, a customer may desire a particular package of options, colors, etc. not currently offered by the manufacturer, while for computer systems, a customer may desire particular RAM, storage, processors, installed adapter cards, etc.

35   In one alternative, a set of potential customers can be selected according to the portfolios to be synthesized and offered. For example, for insurance services relating to

households, potential customers can be homeowners residing in a region of defined insurance risk, such as a particular neighborhood of a city. For goods, potential customers can be identified as past purchasers of similar goods from a certain supplier or in general, or those likely to purchase such goods based on past purchases of related goods. For financial

5   instruments, potential customers can bo those with a certain range of income.

In another alternative, potential customers can make themselves known to a service offering to assemble such custom portfolios of goods, or of services, or of financial instruments. Such services are advantageously specializing according to type of customization provided, and can acquire data, customize portfolios and then offer the

10   customized portfolios using "e-business" methods over the Internet. Alternatively, traditional business methods can be used.

In more detail, this data can be gathered in numerous ways well known to one of average skill in the arts. It can be directly gathered by querying the customers, for example by obtaining responses to surveys and questionnaires, presented, for example, on user

15   devices attached to the Internet. It can be gathered from historical information on the economic or other behaviors of certain potential customers, including, for example, past purchasing choices or investment decisions. This historical information can be known and available to, and provided by, the particular customer or business, or it can be present in databases of economic behaviors from which it is extracted by known "data-mining"

20   techniques. Such economic behavior databases can include data for a single customer, a single store or world wide web site, or for multiple geographically-related or content-related stores or web sites, or can be for even larger economic groupings. In all cases, it is preferable that data gathering, for example, questionnaires, be informed by tools developed in the social sciences, the sciences of opinion sampling, and the economic sciences,

25   particular econometrics.

Regardless of how gathered, such customer preference data can be qualitative, for example, simply an unordered list of desired goods to be insured, desired features of a particular good, types of financial instruments, etc. The data can also be semi-quantitative, wherein also provided, for example, is a relative ranking of portfolio members or sets of

30   members, or relative quantities desired, or so forth. Also, the data can be quantitative with, for example, numerical ratings of preferences which could be target price and quantity ranges.

Having gathered such customer preference data, the methods of the present invention are based, *inter alia*, on the discovery that from such preference data clusters of

35   customers can be discerned that have similar preferences. In a preferred embodiment, customer clusters are discerned by an adaptive dissimilarity partitioning method that

provides both clusters of customer and substantially optimal metrics which measure customer similarity and according to which this clustering can be performed. Next, the methods of the present invention synthesize, for each cluster, a portfolio of goods, or of services, or of financial instruments, or so forth from the predetermined universe of goods,

5   services, or instruments that is customized to best reflect the net preferences of the customers in the particular cluster, while at the same time being profitable to offer at a price satisfactory to the cluster. In a preferred embodiment, the portfolio synthesis is achieved by the innovative methods to be described.

In particular, in the preferred embodiment, in order to synthesize a custom portfolio

10   for a cluster of customers, the methods of the present invention first determine indifference, or utility, surfaces representing the preference data of the cluster. These surfaces represent the net customers' preferences for possible candidate portfolios. In one alternative, indifference surfaces represent options for candidate portfolios for which the cluster of customers is indifferent, in the sense that candidate portfolios described by the surface are

15   all on average equally satisfactory to the customers of the cluster. In another alternative, utility surfaces represent utility values for options for candidate portfolios, the utility values representing the preferences on average of the customers of the cluster for portfolio options. These two alternative representations are readily seen by one of average skill in the art to be substantially equivalent. In either case, these indifference or utility surfaces define a fitness

20   landscape for candidate portfolios from which an optimum portfolio can be selected. An optimum portfolio is selected by searching the indifference surface for portfolios of greater value relative to a current portfolio.

This portfolio fitness landscape has varying ruggedness depending on the preference data gathered. Depending on this ruggedness of this landscape, different search strategies

25   are appropriate to search for optimum portfolios. For example, the nearest neighbor searches are more advantageous on smoother preference landscape, while on more rugged landscapes the advantageous searches include long jumps to more distant neighbors. In detail, the methods of this invention perform multiple objective searches, at least one objective being defied by the portfolio fitness landscape, at least another objective being an

30   economic measure, such as cost or profitability, which reflects the incentives of a provider of the goods, or services or instruments. In a preferred embodiment, the methods of this invention seek a Pareto optimum for the multiple objectives.

In addition to the above-described methods, an object of this invention is to provide systems by which these methods can be performed to offer such portfolio services.

35   Preferably, these methods will be implemented by computer systems in an on-line fashion, for example by use of the Internet, according to which data is gathered directly from

customers of businesses or from records of the economic behaviors of customer or business. Portfolios are then assembled by on-line business-to-business interaction, and are finally offered on-line to customers. Alternatively, these methods can be implemented on "back-office" computer systems interfaced to traditional business methods.

5   This invention achieves the above objects because clusters of customers have more economic "weight," that is more assets or more purchasing power, than any of the individual members, and preferably sufficient "weight" that customized portfolio offerings are profitable at prices acceptable to the customers. Further, these customized offerings, because they are optimal for a cluster or similar customers, are considerably more

10   satisfactory than standardized offerings intended for an entire market. Thereby, this invention improves the marketplace by providing offerings of considerably higher utility than heretofore.

In more detail, the present invention includes the following particular aspects. In a first aspect, the present invention includes a method for dynamically synthesizing portfolios

15   comprising a plurality of individual goods, individual services or individual financial instruments, comprising the steps of gathering preference data from a plurality of customers, wherein the preference data is responsive to the preference of each individual for individual goods, services or financial instruments, partitioning the customers into a plurality of clusters of customers according to said preference data, wherein the customers

20   of each cluster have similar preferences, and synthesizing at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments included in each synthesized portfolio are based on the preferences of the customers of the cluster.

In an alternate aspect, the present invention includes a method for dynamically

25   synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, comprising the steps of gathering preference data from a plurality of customers, wherein the preference data is responsive to the preference of each individual for individual goods, services or financial instruments, partitioning the customers into a plurality of clusters of customers according to said preference data, wherein the

30   customers of each cluster have similar preferences, generating at least one indifference surface for each cluster of customers, wherein a point on the indifference surface indicates the preferences of the customers of the cluster for the portfolio represented by the point, and wherein the indifference surface is based on the preferences of the customers in the cluster, and synthesizing at least one portfolio for each of the clusters of customers, wherein the

35   individual goods, individual services, or individual financial instruments included in each synthesized portfolio are based on the preferences of the customers of the cluster, and

wherein said synthesizing further comprises searching the indifference surface for portfolios indicating relatively greater preference.

The present invention also includes a system for dynamically synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, comprising at least one user device for gathering preference data from a plurality of customers, at least one server computer configured by computer instructions to cause the server computer to gather preference data from a plurality of customers, wherein the preference data is responsive to the preference of each individual for individual goods, services or financial instruments, and to partition the customers into a plurality of clusters of customers according to said preference data, wherein the customers of each cluster have similar preferences, and to synthesize at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments included in each synthesized portfolio are based on the preferences of the customers of the cluster, and at least one communications network for communicating between the user devices and the server computers.

In an alternate aspect, the present invention includes a system for dynamically synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, comprising means for gathering preference data from a plurality of customers at user devices, wherein the preference data is responsive to the preference of each individual for individual goods, services or financial instruments, means for partitioning the customers into a plurality of clusters of customers according to said preference data, wherein the customers of each cluster have similar preferences, and means for synthesizing at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments included in each synthesized portfolio are based on the preferences of the customers of the cluster.

In another aspect, the present invention includes a computer readable medium comprising encoded computer instructions for causing a computer to perform a method for dynamically synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, said method comprising gathering preference data from a plurality of customers, wherein the preference data is responsive to the preference of each individual for individual goods, services or financial instruments, partitioning the customers into a plurality of clusters of customers according to said preference data, wherein the customers of each cluster have similar preferences, and synthesizing at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments included in each synthesized portfolio are based on the preferences of the customers of the cluster.

- 6 -

BRIEF DESCRIPTION OF THE DRAWINGS

Other objects, features and advantages of the present invention will become apparent upon perusal of the following detailed description when taken in conjunction with the appended drawing, wherein:

FIG. 1 illustrates portfolio synthesis method 100 of the present invention;

FIG. 2 illustrates adaptive dissimilarity partitioning method 200;

FIG. 3 illustrates method 300 for determining consumer demand;

FIG. 4 illustrates method 400 for optimizing a portfolio; and

FIG. 5 illustrates representative system 500 on which the embodiments of the present invention can be implemented.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

The present invention includes methods and systems which generally dynamically synthesize custom portfolios of goods, services or financial instruments for clusters of customers determined to have similar preferences. In particular and preferred embodiments, the present invention synthesizes custom portfolios of insurance services for clusters of customers. In the following, this invention is generally described in terms of "customers." It will be understood that such "customers" can be individual persons or can be businesses having any form of organization.

FIG. 1 illustrates the overall method of the portfolio synthesis, method 100, according to the present invention. In step 102, the method begins by gathering customer preference data for a predetermined universe of offerings, such as goods, services, or financial instruments, and a predetermined set of customers, for example those seeking particular custom offerings from a broker or dealer. This data, which can be qualitative, semi-quantitative, or even quantitative as described above, is preferably gathered by reliable techniques, such as those well known in the arts of social sciences, opinion surveying or economic, particularly econometrics. Data can be gathered on-line, for example, over the Internet, or "data-mined" from databases of historic data directed to the economic behavior of customers. It is preferred that customer preference data be as suitably quantitative, in view of the intended portfolio, the desired variability of its elements, and expected prices, and reflect both subjective preferences and objective economic behavior. Such data gathering methods and systems are known to those of average skill in the arts, and this step will not describe further.

Next, in step 104, the methods of the present invention partition the customers into a plurality of clusters according to said preference data. Various known partitioning methods can be used, such as those based on predetermined numerical metrics (for example, a

Euclidean metric) or statistical methods. In a preferred embodiment, the present invention uses an evolutionary learning process called an adaptive dissimilarity partitioning method, which provides both a relevant metric measuring the fit of the clusters to the preference data as well as the set of clusters. In step 105, the method generates indifference or utility

5   surfaces for each cluster of customers, which represent in equivalent form the net preference of all customers of a particular cluster (*i.e.*, the net preference or utility of the cluster's customers) for candidate portfolios of one or more items or elements. These surfaces are based on the preference data gathered from customers within the cluster, and are used in the next step to optimize the value of a candidate portfolio. Finally, in step 106, the portfolio

10   synthesis method creates at least one portfolio for each of the clusters of customers which is particularly adapted to that cluster's key preferences while at the same time being profitably offered at an acceptable price. The synthesis method optimizes multiple objectives to arrive at the portfolio, one objective being the utility represented by the previously determined indifference or utility surfaces, another being the profitability or price of the portfolio. For

15   example, in the case of insurance services, the value at risk and the correlation or negative correlation of the elements of the portfolio is computed; or in the case of stock portfolios, whether the price to earnings ratios of certain issues falls within or without of a predetermined range is determined.

In the following, each of these steps of method 100 is described in its preferred

20   embodiment in more detail.

Turning first to the clustering step, the following conventions will be used in its description. The space of customer preferences is described by a set of n m-dimensional data vectors x_i (i=1, ..., n) having components x_{ij} (j=1, ..., m) which may be real variables, binary variables, or other types of variables. Each component represents

25   preferences for a particular good or service, either quantitative or qualitative, and the preferences of a single customer are represented by one vector. It will be apparent to persons of ordinary skill in the art that other models may be used for the space of preferences and the following methods can be immediately adapted to other such models.

The goal of the clustering step is to assign the customer-preference data vectors to

30   clusters in a manner that minimizes some cost function. A prototype vector is preferably associated with each cluster. A cluster is then defined as the set of customer data vectors that are closer, in the sense of the cost function, to the cluster prototype than to any other prototype.

In alternative embodiments, the clustering step can employ known clustering

35   methods, such as the k-means clustering method or the multidimensional scaling method.

For example, in the k-means clustering algorithm, the coordinates of k prototype vectors y_h (h=1, . . . , k) are determined so that the following cost function is minimized.

$$E_k = \sum_{i=1}^{n} \sum_{h=1}^{k} m_{ih} \left\| x_i - y_h \right\|^2, \tag{1}$$

where $m_{ih}=1$ if x_i is assigned to cluster h and $m_{ih} = 0$ otherwise, and $\| \, . \, \|$ is a distance metric, for example the Euclidean distance, in the space of customer data vectors. The k-means clustering algorithm is explained in McQueen, 1995, Some methods for classification and analysis of multi variate observations, Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1 (Le Cam, L. M. & Neyman, J., editors), University of California Press, Berkeley, CA, pp. 281-297. An acceptable clustering solution is given by $\{m_{ih}\}$, where each data vector is assigned to one and only one cluster. In the k-means algorithm, the cluster prototypes are initialized with the first k data vectors. A new data vector x_i, i > k, is assigned to the closest prototype vector $y_{h(i)}$. The prototype is adjusted in response to x_i, or, more precisely, is moved closer to x_i:

$$y_{h(i)} \leftarrow y_{h(i)} + \frac{1}{\sum_{u=1}^{i} m_{uh(i)}} (x_i - y_{h(i)}). \tag{2}$$

The total adjustment of the prototype is normalized to the number of vectors that have already been assigned to that prototype. A randomized version of this algorithm, supplemented with topological constraints on prototypes, is the self-organizing map, an unsupervised neural network. Unsupervised neural networks are explained in T. Kohonen, 1990, The Self-Organizing Map. Herein a predetermined and well-defined distance is available. If only pair-wise (or higher-order) relationships among vector components are available, then the cost function or metric to be minimized is preferably the product of the dissimilarities of data vectors assigned to the same cluster.

In a further alternative embodiment, multidimensional scaling ("MDS") is used to represent multidimensional customer data points in a two-or three-dimensional Euclidian space such that pair-wise distances in the two or three-dimensional representation space closely match pair-wise dissimilarities in multidimensional space. See, e.g., Cox, 1994, Multidimensional Scaling, Chapman & Hall, London, ("Multidimensional Scaling"). A clustering algorithm can be applied to the representation vectors. Let y_i be the vector that represents data vector x_i. Let d_{iu} be the distance between two representation vectors, y_i and

y_u and D_{iu} the given dissimilarity between x_i and x_u. The cost function (also called stress) is typically given by:

$$E = \sum_{i=1}^{n} \sum_{u=1}^{n} w_{iu}(d_{iu} - D_{iu})^2, \qquad (3)$$

where the weights w_{iu} are introduced to normalize the absolute values of the disparities D_{iu}. A common choice for w_{iu} is

$$w_{iu} = \frac{1}{D_{iu} \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} D_{\alpha\beta}}. \qquad (4)$$

The aforementioned clustering algorithm is applied to minimize the cost and choose the proper representation vectors in Euclidean space. Other definitions of stress and algorithms for minimizing stress are described in Multidimensional Scaling.

In both clustering and MDS, the initial dissimilarity measure, metric, or cost function is assumed known. Given this dissimilarity measure, the clustering algorithm provides clusters, whereas MDS provides a low-dimensional representation preserving clustering. The obtained clusters or representations critically depend on the choice of the dissimilarity measure. Such a measure is usually defined on the basis of "intuitive" criteria and relies on the "expertise" of the system designer. The Euclidean distance metric can be used as a default measure. Defining a dissimilarity measure, however, can preferably be automated. Clustering or scaling data, although it is sometimes used for exploratory data analysis, is usually a first "preprocessing" step in a particular task to be performed (compression, understanding, market segmentation, etc.). The performance of clustering or MDS can therefore be measured not only with respect to the cost function or stress to be minimized but also in connection with the task to be performed.

In the preferred embodiment of the clustering step, the appropriate dissimilarity measure is learned, for example, in a supervised manner on a training set, tested on a validation set, and applied to new data. The preferred learning algorithm is an application of the methods of genetic algorithms ("GA"). Genetic algorithms are described, for example, in Goldberg, 1989, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, (Genetic Algorithms in Search, Optimization and Machine Learning).

FIG. 2 illustrates a flow diagram of the preferred adaptive dissimilarity partitioning method 200. After starting and performing any necessary initialization, the method 200

chooses, in step 202, a generic family of distance metrics or dissimilarity measures. In step 204, the method 200 randomly generates a population of dissimilarity measures

$$D^v = \left\{ D_{iu}^v \right\}$$ or distance functions d^v within the chosen generic family, where v is the index

5   of a given dissimilarity measure in that population. The parameters of each "individual" D^v are encoded into a "genotype" according to GA methods. In step 206, the method 200 performs clustering or multidimensional scaling with each generated distance function or dissimilarity measure. In step 208, the method 200 evaluates the performance of clustering or multidimensional scaling and assigns fitness to every dissimilarity measure D^v. In step

10   210, the method 200 selects individual measures on the basis of their fitness. In step 212, the method 200 applies known operators to the "genotypes" of selected individual measures and selected pairs of individual measures. Preferably, the operations are known genetic operators, such as mutation and crossover.

In step 214, the method 200 determines whether the partitioning results are

15   satisfactory with respect to the fitness computed in step 208. If the partitioning results are not satisfactory, control returns to step 206 to perform clustering or multidimensional scaling for each new distance function or dissimilarity measure created in steps 210 and 212. If the partitioning results are satisfactory, control proceeds to step 216 where the method 200 terminates.

20   The distance function or dissimilarity measure can be represented by a true function of the vectors coordinates or by a set of pair-wise relationships. When only pair-wise relationships between data vectors are available, generalization of the dissimilarity measure to data vectors which have not been represented (using, for example, MDS) is needed. The simplest generalization procedure is to use a locally linear interpolation, using the k nearest

25   neighbors: the dissimilarity between the new vector V and any other vector W is given by the average dissimilarity between the k nearest neighbors of V and W.

The following example illustrates the operation of the adaptive dissimilarity partitioning method 200. Let us assume for definiteness that each data vector x_i is two-dimensional. The two components of x_i represent, for example, two properties of a

30   mortgagee providing mortgage services, for example, level of customer service and relative cost to refinance, on a scale of one to ten. A set of n customers is asked to determine the level of customer service and the relative cost to refinance that they desire in their mortgagee. In addition, each customer is asked to tell who the mortgagee is. Assume that k different types of mortgagees are represented. The distance function in the space of

35   customer preferences is unknown. For example, one factor may be more important than another. A simple family of distance functions is:

$$d_{iu} = \left| f_1(x_{i1}, x_{i2}, x_{u1}, x_{u2})(x_{i1} - x_{u1})^2 + f_2(x_{i1}, x_{i2}, x_{u1}, x_{u2})(x_{i2} - x_{u2})^2 \right|^{1/2}, (5)$$

where f_1 and f_2 are, for example, second-degree polynomial functions of their variables. Each function is characterized by 15 parameters, the coefficients of the polynomials. The variation of these parameters is assumed to be restricted to [-10,10]. A clustering algorithm, such as k-means, is applied to the data set using this distance function. The fitness of a distance function d^v is given by:

$$F^v = \frac{1}{1 + M_{in} + M_{out}}, \qquad (6)$$

where M_{in} is the number of customers assigned to the same cluster that do not use the same mortgagee and M_{out} is the number of customers assigned to different clusters that use the same mortgagee. Depending on the task at hand, these two types of mismatches can be given different weights.

The best individuals obtained after, one thousand generations of the genetic algorithm, corresponded to distance functions that produce the clusters of customers with the most favorable fitness described above.

The adaptive dissimilarity partitioning method 200 finds the natural dissimilarity measure or distance function in a space of attributes. This function may be unknown. Instead of resorting to *ad hoc* functions, the method systematically generates a distance function adapted to the task at hand. The obtained distance function reflects the structure of the space of attributes and therefore can be used to cluster customers, extract the "natural" clusters in the data using a non-parametric clustering algorithm (that is, one in which in the number of clusters is not predefined), and extract the effective dimension of the space of preferences.

In another example, two hundred two-dimensional data vectors were randomly generated. Let x_{i1} and x_{i2} be the x- and y-coordinates of the ith data vector. x_{i1} and x_{i2} are drawn from a uniform random distribution on [0,1]. Let us assume that x_{i1} and x_{i2} represent customer preferences for two selected features of a given product type, that two products are on the market, and that customer i purchases product one if and only if $x_{i1} < 0.5$ and purchases product two if and only if $x_{i1} \geq 0.5$. In this example, therefore, only x_{i1} is relevant in the determination of what product is purchased by a customer whose preference vector is (x_{i1}, x_{i2}). But this information is not known to the analyst, who simply assumes that the relevant distance in preference space is, for example, the Euclidian distance. Using such a distance, the analyst will be unable to correctly segregate customers into two classes. What

the algorithm has to find is the relevant distance in preference space that will naturally lead to the correct segregation after application of a simple clustering algorithm. Here we use a modified version of the k-means clustering algorithm with k=2. Two centroids are initially located at (0.5, 0.25) and (0.5, 0.75). After application of the clustering algorithm with the

5   appropriate distance function, the centroids should converge to (0.25, 0.5) and (0.75, 0.5), which best represent the purchase/not-purchase decision clusters. With this clustering algorithm a data vector belongs to the cluster whose centroid is closest to that data vector.

Let $C_{m(i)}$ be the centroid closest to vector x_i ($C_{m(i)} = ArgMin\left(d\left(C_m, x_i\right)\right)$, where d is the

10   distance function), and $C_{m(i)j}$ the jth coordinate (j=1,2) of $C_{m(i)}$. The centroid update function upon presentation of the next data vector, x_i, is given by:

$$C_{m(i)j} \leftarrow C_{m(i)j} + \eta \frac{d(C_{m(i)}, x_i)}{n} \sigma(x_{ij} - C_{m(i)j}), \qquad (7)$$

15   where d is the current distance function, $\sigma(\ )$ is the sign function ($\sigma(u)=+1$ if $u>0$, $\sigma(u)=-1$ if $u<0$, and $\sigma=0$ if u=0), η is a learning rate, and n=200 is the number of data vectors. The family of distance function used in this example has three parameters:

$$d\left(x_i, x_h\right) = \left[w\left|x_{il} - x_{j1}\right|^\alpha + (2-w)\left|x_{i2} = x_{j2}\right|^\beta\right]^{\frac{2}{\alpha+\beta}}, \qquad (8)$$

20

where w, α, and β $\in[0,2]$. When w=1 and $\alpha = \beta = 2$, the usual Euclidian distance is recovered, and when w=1 and $\alpha = \beta = 1$, then the function becomes the city-block (or L_1) distance.

This family of distance functions can easily be generalized to higher-dimensional

25   spaces. For example, let us consider a D-dimensional space:

$$d\left(x_i, x_h\right) = \left[\sum_{p=1}^{D} w_p \left|x_{ip} - x_{jp}\right|^{\alpha_p}\right]^{\frac{D}{\sum_{p=1}^{D} \alpha_p}}, \qquad (9)$$

30   with

$$\sum_{p=1}^{D} w_p = D, \qquad (10)$$

35   where α_p(p=1, . . . ,D) and w_p (p=1, . . . ,D) are 2D parameters (of which only 2D-1 are free parameters) that determine the relative importance of the pth coordinate and the amount of

distortion along the pth coordinate. This family of functions assumes no correlation among coordinates. When such correlation is present, other distance functions should be used in such cases, for example, with cross terms in the coordinates.

For the two-dimensional example, a fitness-proportionate genetic algorithm ("GA") was used with the following fitness function for distance D^v:

$$F^v = \frac{1}{1 + M_{in} + M_{out}},\qquad(11)$$

where M_{in} is the number of customers assigned to the same cluster that do not purchase the same product and M_{out} is the number of customers assigned to different clusters that buy the same product. The GA parameters are as follows: the population size was forty; the mutation rate was 0.1; and the crossover operator was replaced with averaging of parameters (that is, two selected individuals produce one offspring, the parameters are the arithmetic average of its parents' parameters).

After 10 generations, the GA found values of the parameters that consistently produce a perfect clustering of customers after application of the modified k-means algorithm. In contrast, during one application (200 iterations) of the k-means algorithm alone (without GA learning of the distance metric) for initially "bad" values of the parameters ($w=0.96$, $\alpha=1.81$, $\beta=1.77$), close to the Euclidian distance, the centroids are unable to move to the optimal locations and remain confined in the vicinity of their initial values. For initially "good" values of the parameters, as found by the GA after 10 generations ($w=1.98$, $\alpha=1.67$, $\beta=0.03$), the centroids moved to the optimal locations because the distance function assigns almost all the weight to the x-coordinate. The GA has therefore been able to find a good distance function, from within the family of distance functions, that reflects the structure of this exemplary preference space.

Assume now that instead of being uniformly distributed in [0,1] x [0,1] customers form four clusters (with the same "purchase" rule: a customer i purchases product one if and only if $x_{i1}<0.5$ and purchases product two if and only if $x_{i1} > 0.5$). Two situations can occur: the four clusters may discriminate along the y-axis or along the x-axis. Upon application of a non-parametric (an undefined number of clusters) clustering or multidimensional scaling algorithms, the situation where the four clusters may discriminate along the y-axis should lead to the detection of two clusters while the situation where the four clusters discriminate along the x-axis should lead to the discovery of four clusters if the appropriate distance function is used. If the Euclidian distance function is used both situations lead to the detection of four clusters. A non-parametric algorithm leads to four clusters in both cases using the Euclidian distance. The same algorithm leads to two clusters when applied to the

situation where the four clusters discriminate along the y-axis and four clusters in the situation where the four clusters discriminate along the x-axis.

In an alternate embodiment, for more complicated problems, general function approximators can be used. An example of a known general function approximator is
5   neural networks. In the case of neural networks, the connection weights are evolved using the genetic algorithm as described above.

In another alternate embodiment, the GA is interactive: the outcome of the clustering or MDS algorithm is evaluated by a human observer who picks the good solutions, *i.e.*, the observer assigns the fitness.

10   Next, the methods of the present invention, which determine portfolios satisfying consumer preferences, determine the context dependent, combinatorially optimized set of properties, uses, or features that are important for optimizing for customers the value of portfolios of goods, services, or financial instruments. The properties, uses or features are determined by computing and examining a plurality of indifference, or equivalently utility,
15   surfaces for each cluster of customers.

Parameter models of indifference or utility surfaces are known. Preferably, the indifference surfaces of this invention are modeled by such parameterized models. A preferred model is the NK model described in Stuart A. Kauffman, 1993, The Origins of Order, Oxford University Press, Chapter 2, and in Stuart A. Kauffman, 1995, At Home in
20   the Universe, Oxford University Press, Chapter 9. The "ruggedness of NK models of fitness landscapes is parameterized by K, the larger is K (K is always less than N) the more rugged the landscape is. A landscape is called rugged if, intuitively, there are many local peaks, or maxima, of many sizes at many spacings, or equivalently, if the landscape correlation falls off rapidly with increasing separation distance. Conversely, a correlated
25   landscape with a few well-positioned peaks is called smooth.

NK landscapes are members of a still more general class of models in physics, known in the art as order-P spin-glass models. An order-P spin-glass model consists of N spins, each of which can take on a discrete number of values, e.g. -1 and +1, or 1 and 0, or a, b, c, d. Each spin contributes an "energy" to the total energy of a system of N spins. The
30   energy of a given spin configuration of the N spins is given by the sum of the energies of the N spins. Each spin's energy contribution is, in general, given by a sum of a monomial term which is a function of its own state, plus quadratic terms which are sums of energies that are functions of the states of all spins that influence it in pair-wise interactions, plus a similar sum of cubic terms listing all the contributions of all triples of spins of which that
35   spin is a member, plus higher order terms up to order P. In the NK model, K is the highest order coupling.

In such spin-glass models, the discrete system has rugged "fitness," "cost," "efficiency," or "utility" landscape over the combinations of states of the N spins. Techniques have been developed to characterize a number of features of such landscapes. And these features allow ready assessment of the importance of higher order, combinatorial

5   properties of landscape structure. The properties include five features: 1) the number of peaks in the landscape; 2) the expected number of steps to a peak from any-given point in the landscape; 3) the rate of decrease dwindling number of directions "uphill" (in directions of increasing fitness or utilities) as a peak is climbed; 4) the number of different peaks that can be climbed from a single point on the landscape by adaptive walks proceeding only

10   uphill; 5) the correlation structure of the landscape which is, measured by the correlation between fitness at two points on the landscape as a function of their distance. According to this invention, therefore, such parameters are derived from customer preference data in order to characterize the landscape of customer preferences or utilities.

These properties of discrete landscapes, where the spins take on only discrete values,

15   a, b, c, d... are generalized in the case of continuous dimensions, where each variable is a real number. In this continuous case, the lengths of walks uphill, and dwindling directions uphill are parameters of a "step length." In a space of reasonably smooth hill sides, any point on the landscape that is on a hillside has the property that, for infinitesimal steps away from that point, half the directions are uphill and half are downhill. Only on ridges, saddles

20   and peaks is that false. However, if a discrete step length, e.g., 100 yards, is specified, then as a path continues uphill and a ridge or saddle or peak is approached, the "cone" of directions that are still uphill will decrease. The rate of decrease is another measure that can be used to characterize the ruggedness of a continuous landscape. Thus, on NK landscapes, with K modestly large (for example, K>5), the generic feature is that at every step uphill,

25   the number of directions uphill falls by a constant fraction. As landscape ruggedness increases, the fraction by which the direction uphill dwindles increases from a few percent to 50% for fully random landscapes in the K = N-1 "random energy" limit. In a similar way, the rate at which the uphill cone of directions decreases as walks uphill continue provides a further measure of landscape ruggedness for continuous landscapes.

30   Now in detail, consider the universe of goods, services, or financial instruments out of which the present invention synthesizes optimal portfolios. Without loss of generality, the following description is that of mortgage services offered by a mortgagee. Other application of the present invention, such as to goods or to insurance services to stock portfolios, will be immediately apparent to one of average skill in the art. Certain

35   preferences important to mortgages were described above; many other preferences are widely known. Consider, to be concrete and without loss of generality, discrete choice

data-gathering methods. A customer is presented with different choices of a bundle of properties, or vector of properties. Each bundle is a point in the property space. A price is attached to each such point, *i.e.*, the closing costs. The customer is asked to choose which, if any, products would be just acceptable. Examination of the vectors in the property space

5   found after a several such choices, determines a cost, such that in the vicinity of those positions (indifference points) in property space having this cost the customer will just stop choosing. Thus, further, from such points a surface can be found in property space having this cost, such that on one side of this surface, the customer will not choose while on the other side of this surface, the customer will choose. This surface estimates the price for that

10 specific vector of properties. By sampling at many points for one customer, it is possible to build up this utility surface in property space at one cost for that customer, equivalently an indifference surface. Further data gathered for different prices builds up a set of such surfaces at the different prices.

Similarly, by considering all the customers in the cluster, a population of such

15 indifference data points can be determined, and from such data points, a set of indifference surfaces at various prices can also be determined for all the customers in the cluster. The input to this determination, the customers' preference data gathered at step 102, preferably is gathered according to the following criteria: first, this data is obtained over a moderate large region (at least one quarter, preferably at least one half) of property space. The data

20 points are then typically each labeled by a vector of preferences, and, using standard analysis, both high utility positions in the space of properties are discriminated in order to optimize the vector of goods produced, each at a different position in the property space.

Preferably according to this invention, parameters reflecting landscape roughness are used to improve and focus the above standard procedures employed for data gathering.

25 These parameters direct limited sampling to capture higher order landscape structure through determination of the context dependent (that is local) features of these landscapes. Landscape parameters also help build statistical models of an "equivalence class" of the real landscape, and can also be utilized to build actual models of the actual market scape.

Fig. 3 illustrates a flow diagram of preferred method 300 for determining

30 indifference, or utility, surfaces that find the context dependent, or combinatorial optimized set of properties, uses, or features (for example, landscape parameters) that allow optimization of the value of portfolios products to the customer cluster. In step 302, method 300 selects an indifference point in property space that lies on a surface that divides a region of product portfolios where a predetermined customer would buy from a region of product

35 portfolios where the predetermined customer would not buy.

In step 304, the method samples in a determined and directed manner a set of points on a R-dimensional sphere surrounding the point selected in step 302. Step 304 contrasts with known methods for predicting consumer demand that sample widely and uniformly over product space. In the method of the invention, the radius of the sphere is defined as the

5   "step length" on the indifference surface, which is chosen according to surface ruggedness, or equivalently, landscape parameters, in order to determine efficiently significant structure of the surface. An exemplary distance is the Euclidian distance. With the same customer, or more generally, the same cluster of customers, step 304 characterizes for many points in the spherical surface surrounding the point whose price has been determined, whether that

10   new point would or would not be purchased by the customers of the cluster at the given price. Since the true price surface in the space of properties contains the first determined point, that price surface will, in general, pierce the spherical surface surrounding the point whose price is determined. The points on the sphere which are purchased and the points which are not purchased determine a curve of points marking the transition between buying

15   and not buying at the price. In this way, the neighboring indifference surface hood surrounding the first indifference point can be determined.

In step 306, method 300 determines whether the indifference surface has been substantially completed. If the indifference surface has not been substantially completed (for example, covering at least one half of the possible portfolios), control proceeds to step

20   308. In step 308, the method selects another point on the indifference surface from the transition curve determined in step 304. After step 308, control returns to step 304. Step 304 samples a set of points on a R-dimensional sphere surrounding the point selected in step 308. In this fashion, method operates to extend the indifference surface at the predetermined price through the property space of possible portfolios.

25   The ruggedness of the indifference surface at a given price is reflected in the previously-discussed parameters. Thus, measured in property space, the indifference surface at a given price can have one or more correlation lengths. These correlation lengths, in the NK (the order of coupling in this model is K) model are long, for K small (a smooth surface), and short for K large (a rugged surface). Thus, short correlation lengths are due to

30   and estimate higher order couplings among portfolio contents. The cone of "uphill" directions in property space on an indifference surface at a given price can be determined. Good combinations of properties will show up as peaks or minima, depending upon direction of definition, in the surface. That is, a good combination of properties in property space will show up, for example, as a willingness to pay the fixed price for a small

35   "amount" of the given vector of properties. Having defined a local "peak" in the indifference landscape surface, an optimum walk length, step size, peak, and number of

peaks to which one can walk from any point. These parameters are used to control searches for an optimum portfolio. In addition, the similarity of peaks climbed from the same or nearby points on the indifference landscape at a given price can be examined. Accordingly, it can be determined if high peaks cluster near one another, recombination (used in a GA) is

5   a good means to search for high peaks. From answers to such question, a search for high peaks can focus in precise ways to search between high peaks on the current landscape, and hill climb from those points to still higher peaks.

Thus, determination of landscape properties and parameters enable focused sampling during the data gathering steps of the landscape to estimate the higher order

10   context dependent, combinatorial features of a given market scape.

Alternatively, statistical models of the sampled market scape can also be built by utilizing order-P spin-glass-like models, where the class of models with all possible values of the coefficients of all the P-adic terms in the polynomials constitutes the family of landscape models. Maximum entropy Bayesian updating techniques can then be used to

15   estimate the most likely landscape parameters to fit the observed data.

A major improvement of the present invention and known methods is that the detailed sampling in specific regions of the indifference surface at a given price yields estimates of the how "high" the higher order terms, (K in the NK model) actually are. Thus, from such focused local measurements at several points on the landscape, it can be

20   determined that, for example, fifth order interaction, P=5, are critical for determining the local structure of the market scape. Knowing that, a preponderance of the data can be gathered and used to fit or estimate the 5th order term, while only a small amount of data is gathered and used to estimate the monomial terms (that determines the overall non-isotropic features of the market scape on long length scales across the market scape). Thus, data

25   gathering can be optimized to discover both long range features of the landscape and local features.

Given this analysis, one can derive a class of statistical models of the landscape, and specific models of the landscape which are preferably parameterized by parameters of landscape ruggedness or smoothness.

30   Step 105 was explained above in the context of computing an indifference surface for a predetermined price in the property space of mortgage services for a predetermined customer. However, as will be known by one of ordinary skill in the art, method 300 can also be used to sample the property space of the product for a given cluster of customers at a predetermined price or at a set of predetermined prices. This procedure defines one or more

35   optimal customer features for a given mix of goods (or services or investment instruments) or position, in product space. The same procedure allows multiple points in product space

to be utilized, indeed just the points normally utilized, to find the best set of positions in product space to match the best targeted populations of customers in customer preference space. Again, the advantage of present invention is that it allows the higher order terms, the context dependent features in customer preference space, to be more readily detected, for it

5 tells us that K order terms are important. Again, statistical models of customer preference scapes, and models of specific customer preference scapes, can then be constructed.

Next, method 100 of the present invention, at step 106, synthesizes a portfolio of goods, services or financial instruments which is optimized to fit the preferences of each cluster of customers. In general, in this step, landscapes of various types, in particular

10 indifference surfaces of customer preferences previously determined, are searched for optima, which may be maxima or minima depending on the landscape type. According to this invention, such searches are preferably performed by starting from an initial portfolio and examining neighboring portfolios for increased fitness. The distance to neighboring portfolios, their direction, and other selection parameters are selected according to the

15 landscape parameters determining ruggedness or smoothness. If an improved portfolio is found, the search is started from that portfolio. The invention is adaptable to other search methods responsive to landscape parameters. For example, a genetic algorithm can search by "evolving" a population over a landscape, where the parameters of the "evolution" are chosen according to landscape ruggedness and other landscape parameters.

20   In more detail, in most applications two or more landscapes are simultaneously searched for optima. Usually, at least one landscape is the landscape of the net customer preferences of the customers in a cluster for the components of a candidate portfolio. Another usual landscape is one that represents the feasibility of providing the candidate portfolio. For example, for goods or services, such feasibility can be represented by a

25 landscape determined by an economic (for example, cost) or a technological (for example, manufacturability) function of the candidate portfolio. For financial instruments, the feasibility is an economic landscape responsive to the methods and costs of acquisition or divestiture of the particular instruments of interest. For insurance services, the feasibility is also economic and is a function of, for example, the historic risk of loss for the goods in the

30 portfolio in the geographic locations of the customers.

In a preferred embodiment, optimization of multiple objectives is performed in order to reach a Pareto optimum. A Pareto optimum portfolio for multiple objectives is one such that any possible portfolio change will reduce the fitness of at least one objective even if the fitness of another objective is increased. Therefore, by combining multiple objectives

35 according to a Pareto ranking or ordering, multiple objectives can be optimized for a portfolio in a manner substantially identical to optimizing a single objective for a portfolio.

In the following, two alternatives are described for optimizing a single objective, which therefore are immediately applicable in general to step 106 of method 100. The first alternative is described in terms of minimizing a value at risk (VaR). This alternative is particularly preferable to, for example, portfolios of financial instruments or of insurance

5 services. For insurance services, historical statistical records on the risk of loss of possible assets to be insured is input to optimizing a portfolio by limiting the VaR. For financial instruments, historical records of past transactions involving possible instruments is input. The following, without loss of generality, is directed primarily to the case of financial instruments, particularly publicly traded stocks or bonds.

10 In detail, an initial portfolio can be generated from available historical data. The historical simulation method of the present alternative generates an initial portfolio of products based on historical data to minimize the value at risk of the portfolio. If historical data is not available, an initial portfolio can be generated consisting of the entire market of available products, and the remainder of this method can be skipped. Value at risk is a

15 single, summary, statistical measure of possible portfolio losses. Alternatively in the goods context, one could substitute the total cost of the portfolio for value at risk, recognizing that some suppliers will discount cost for large orders of several different goods. Specifically, value at risk is a measure of losses due to "normal" market movements. Losses greater than the value at risk are suffered only with a specified small probability.

20 Using a probability of x percent and a holding period of t days, a portfolio's value at risk is the loss that is expected to be exceeded with a probability of only x-percent during the next t-day holding period.

The technique to minimize the value at risk utilizes historical simulation. Historical simulation requires relatively few assumptions about the statistical distributions of the

25 underlying market factors. In essence, the approach involves using historical changes in market rates and prices to construct a distribution of potential future portfolio profits and losses, and then determining the value at risk as the loss that is exceeded only x percent of the time.

The distribution of profits and losses is constructed by taking a current initial

30 portfolio, and subjecting it to the actual changes in the market factors experienced during each of the last N periods. That is, N sets of hypothetical market factors are constructed using their current values and the changes experienced during the last N periods. Using these hypothetical values of market factors, N hypothetical mark-to-market portfolio values are computed. From this, it is possible to compute N hypothetical mark-to-market profits

35 and losses on the portfolio.

The following discussion describes the technique for isolating low value at risk portfolios. Consider a single instrument portfolio, in this case stocks traded on the New York Stock Exchange and NASDAQ markets. For this instrument, there exists tremendous amounts of data. If we assume a one day time horizon (t= 1), then the data we are interested

5   in are the daily closing prices of every publicly traded stock on the two markets. Such data exists for thousands of stocks for tens of thousands of days. From these data, it is possible to construct an m x n matrix (where m is the number of stocks, and n is the number of days) of prices.

Within this collection of stocks, there are pairs, triplets, quadruplets, etc., of stocks

10 whose values at risk are lower as a group than any of the stocks individually. This occurs because sets of stocks whose price changes are anti-correlated will have lower values at risk than the stocks individually. When the price of one stock goes down, the price of the other tends to go up. The chance that both stocks go down together is lower than the chance that two stocks chosen at random would go down together because the stocks are anti-correlated.

15 This reduces value at risk.

The optimal portfolio would group anti-correlated stocks in the optimal proportions to minimize value at risk. Because there are so many stocks, however, the space of all possible portfolios is too large to search exhaustively. Genetic algorithms are well suited to finding good solutions to this problem in reasonable amounts of time. The algorithm works

20 as follows:

Step 1:

Start with m portfolios. Each portfolio can be represented as a vector of length m. Each bit (m_i) in the vector is either a 1 or a 0 signifying that the ith stock is either

25   included or excluded from the portfolio. This can later be extended to letting each bit specify the number of shares held rather than simply inclusion or exclusion. To each portfolio, assign a random number of stocks to hold such that every possible portfolio size is covered (at least one portfolio excludes all but one stock, at least one portfolio excludes all but two stocks, and so forth, and at least one portfolio

30   includes all the stocks). Once the number of stocks to hold has been assigned, let each portfolio randomly pick stocks until it has reached its quota.

Step 2:

Go back in time n/2 days (halfway through the unexamined data). For each of the m

35   portfolios, compute the value at risk for the n/2 days that precede the halfway point.

Step 3:

Randomly pair portfolios. For each pair of portfolios, let the portfolio with the higher value at risk copy half of the bits of the lower value at risk portfolio (i.e. randomly select half of the bits in the more successful portfolio. If a bit is different, the less successful portfolio changes its bit to match the more successful portfolio). The portfolio with the lower value at risk remains unchanged.

Step 4:

Repeat steps 2 and 3 for the unexamined half of the data (replacing the number of days, n, with n/2) until a threshold for value at risk is achieved.

In this way, clusters of anti-correlated stocks spread through the population of portfolios. This method ultimately selects for most or all of the good clusters. Notice that this method may also alight upon the optimal number of stocks to hold in a portfolio. For example, if the minimum VAR portfolio contains only three stocks, three-stock portfolios will tend to propagate through the population.

Finally, the present invention optimizes the initial portfolio generated by using a method of sampling and selection to evaluate and minimize risk for a portfolio of assets with uncertain returns, while at the same time maximizing any of the optimal customer features identified by the indifference surface analysis. The present invention involves risk management techniques which move beyond pair-wise value at risk and risk analysis in general, which optimize any figures of merit, including customer preference. This extension will be described after considering the case where risk is the sole feature to be evaluated.

In risk analysis where the future rewards are uncertain, there are two important concerns of the holder of the portfolio. First, it is important to quantify the risk (the amount of money that could be lost) over some time horizon. Second, the holder wishes to structure the portfolio so as to minimize the risk.

Let $x_i(t)$ represent the value at time t of the ith asset in the portfolio. If there are N assets in the portfolio let $x(t)$ be the N-vector representing the values at time of all components of the entire portfolio. The value of the entire portfolio to the holder is specified as some function f(x) of the values of the assets. Typically, this function might be a linear combination of the asset volumes times the asset prices. i.e. $f(x) = \sum_{i=1}^{N} v_i x_i$.

Furthermore let $P(x', t'|x, t)$ represent the probability that the asset prices are x' at time t' > t given that the asset prices were x at time t. If t indicates the present time and x represents

the present value of the assets then the expected value of the portfolio at some time t' in the future is:

$$V(t'|x,t) = \int dx' f(x') P(x',t'|x,t).$$ (12)

This value indicates the expected worth of the portfolio but does not reveal what the risk is, *i.e.* what might conceivably be lost. To determine this quantity, from $P(x',t'|x,t)$ can also be determined the probability $P(v|t)$ that the value at time t is v:

$$P(v|t) = \int dx' \sigma(v - f(x')) P(x'|x,t).$$ (13)

This probability is the fundamental quantity which allows assessment of risk since it gives the probabilities for all potential outcomes. Thus, for example, statement like "with 95% confidence the most money that will be lost, is v^*" can be made. In this case v^* is determined from the requirement that only 5% of the time will more money be lost, *i.e.*

$$\int dv \ P(v|t) = 0.05.$$ (14)

Other measures of risk are similarly based on $P(v|t)$.

The risk will depend sensitively on the precise form of $P(x',t'|x,t)$. Consider a pair of assets i and j that are anti-correlated with each other *(i.e.* when the price x_i increases the price x_j usually decreases). If one invests equally in both assets then the risk will be small since if the value of one asset goes down the other compensates by going up. On the other hand if the price movements of assets are strongly correlated then risks are amplified. To evaluate and manage risk it then becomes paramount to identify set of assets that are correlated/anti-correlated with each other. This observation forms the basis of traditional value at risk analyses ("VAR") in which the risk is assessed in terms of the covariance matrix in asset prices. The covariance matrix includes all the possible pair-wise correlations between assets.

While traditional VAR captures pair-wise variations in asset prices it completely ignores higher order relationships between variables, e.g. when assets t and j go up asset k goes down. Moreover the Gaussian assumption inherent in VAR is known to be false. What is needed is a more general approach. The present invention includes new risk management

techniques which move beyond pair-wise VAR. The preferred embodiment utilizes schemes to accomplish higher ordered VAR.

One method which recognizes that information about higher order relationships can be uncovered by looking at the VAR of subsets of assets from the portfolio is called cluster

5   identification. Consider that a specific set of assets covaries with each other in some predictable way. Knowledge of this covariation can be used to devise a risk adverse combination of these particular assets. Since the variation involves all four assets it can never be determined by only looking at pairs of assets. Noting that the historical record of asset prices and portfolio values provides a training set, clusters can be discovered from this

10 set. The historical record provides a data set which includes the true VAR, because the future value of the portfolio is known from the historical data. Let v represent the true VAR for a particular portfolio x at a point T into the future. From the historical record, form the data set $D = \{x_i, v_i\}$ and thus estimate the VAR for the assets in the chosen portfolio, *i.e.*

15 $P(v|x)$. If one assumes that the stochastic process that generated D is stationary then the same relationship discovered in D will also hold in the future. Once the mapping from a cluster set to a VAR has been determined, search over the subsets to find a combination that gives particularly low VAR.

Begin by making the simple assumption that $P(v|x) = \delta(v-\mu(x))$, *i.e.*, it is

20 characterized entirely by its mean value $\mu(x)$. This mean value will differ for different subsets of assets. In a more elaborate embodiment, the variance around this mean could also be included, using an assumed Gaussian distribution of fluctuations: $P(v|x) = N(\mu(x), \sigma^2(x))$. From the data D, much more complicated relationships could be inferred, but, without limitation, the present discussion is in terms of this case.

25 Given that one can determine the true average VAR for any set, identify those assets within a portfolio of N assets that form good combinations. Computationally the following scheme can be used to identify good subsets of assets. Assume that the optimal subset of assets is of size n « N. Starting from the original portfolio randomly form portfolios of half the size by sampling (without replacement) from the entire portfolio. The

30 probability that any one of these randomly generated portfolios contains all n assets is approximately $1/2^n$. Thus, in significantly more random portfolios than this it is likely one will obtain at least one subset containing all assets. For each of the randomly generated portfolios of N/2 assets, determine its VAR by calculating it from D and keep those portfolios with high VAR. In this way, only the most promising portfolios, *i.e.* those that

35 contain the subset sought, are kept. This process can then be iterated further. From these remaining portfolios of size N/2, randomly generate portfolios of half the size (N/4). Assuming that, at least one of the size N /2 portfolios contained the desired cluster the

probability that one of the size *N/4* portfolios contains the full subset, is again $1/2^n$. Keep iterating this process of generating and filtering portfolios and each time comes closer to good subsets.

After *m* iterations of this procedure the portfolio size is $N/2^m$. Let $\overline{m}$ be the largest

5 value of *m* such that $N/2^m$ *is greater (i.e.* the largest portfolio that contain all *n* assets) and let $\underline{m} = \overline{m} + 1$. An abrupt increase in the VAR from $\overline{m}$ to $\underline{m}$ since will occur, since form a risk adverse combination of all the *n* assets at $\underline{m}$ can no longer be formed. This fact indicates that *n* must lie between $N/2^{\overline{m}}$ and $N/2^{\underline{m}}$. At this point, samples from the portfolio of size $N/2^{\overline{m}}$ can be taken to form new portfolios of size

10 $(N/2^{\underline{m}} + N/2^{\overline{m}})/2$. The extreme VAR values of these new portfolios will be

comparably to either the $N/2^{\overline{m}}$ in which case $(N/2^{\overline{m}} + N/2^{\underline{m}})/2 \le n \le N/2^{\overline{m}})$ or

comparable to $N/2^{\underline{m}}$ in which case $N/2^{\underline{m}} < n < (N/2^{\overline{m}} + n/2^{\underline{m}})/2$. Iterating this

15 procedure determines the optimal subset size *n*. Knowing the optimal *n*, different subsets of this size are searched to eventually pick out the precise combination of the *n* assets.

There are many variations to this basic method that can improve efficiency. The portfolio size can be reduced by a fraction other than one half at each step, since a higher probability of retaining the subset intact is sought. The best number of random portfolios to

20 generate and test can also be adjusted to make the search more efficient. Simple analytical model can be built to optimize these algorithm parameters.

As previously described, this above method to minimize VAR can be extended to determine subsets with other desired properties, with respect to other objectives, including those identified by customer-preference indifference surfaces. For example, suppose that in

25 addition to risk aversion, the indifference surfaces are used to identify that providers of portfolios also wanted to maximize profit. Also the customer clusters might seek to balance risk/reward. To extend the above method to handle multiple objectives, sub-sampled portfolios are generated but the selection criteria amongst portfolios is modified. Instead of picking sub-sampled portfolios which have the best VARs we measure, a number of

30 objectives for each of the particular sub-sampled portfolios are evaluated and those sub-sampled portfolios which Pareto dominate all other portfolios (generated at the present iteration or all previously generated portfolios) are kept. Except for this selection criteria change, the remainder of the above method is unchanged. Upon termination, a portfolio which is Pareto dominant with respect to all objectives (for example, specified by

35 indifference surfaces) is obtained.

The present invention also includes a method for portfolio modification. There are other methods to try to identify beneficial changes to a portfolio. Traditional VAR theory measures the effects of modifying (i.e. increasing or decreasing the holding) a position in one of the assets. As seen earlier, if higher order combinations of assets are important then

5   the effects of a single asset might be minor. There is an important practical reason why traditional VAR focuses on the changes of only a single asset. If the portfolio is on size N and we consider changes involving m assets then on the order of N^m stocks must be examined. Consequently, for practical reasons attention is restricted to m=1 or single asset changes.

10   A second alternative for optimizing a single objective, which is also immediately applicable in general to step 106 of method 100, is method 400 illustrated in Fig. 4. This alternative can be used also to determine the optimal number of assets to change while searching for an optimal portfolio. In step 410, the method inputs or determines the landscape parameters of the fitness landscape of an objective to be optimized. For example,

15 the landscape can be defined over the portfolios as the preferences of clusters of customers, or over the portfolios of financial instruments as VAR described above. For a further example, the landscape can be modeled with an NK model, and the parameters input are N, K and the functions necessary to define the dependence of the fitness on the K neighbors. Two portfolios are neighbors if they differ in the holding of a single asset. Alternatively,

20 the landscape can be inferred from historical data using techniques described in the co-pending application titled, "An Adaptive and Reliable System and Method for Operations Management," U.S. Application No. 09/345,411, filed July 1, 1999.

In step 420, the method determines a substantially optimal searching distance, d*, by processes described in co-pending international application designating the United States

25 No. PCT/US99/19916, titled, "A Method for Optimal Search on a Technology Landscape," filed August 31,1999. This process is responsive to the ruggedness of the fitness landscape, or to parameters modeling this ruggedness.

In a first alternative, the searching distance is determined with the NK model is used to model the fitness landscape. First, a correlation coefficient is derived for the *NK* model

30 landscape. Suppose the portfolio is changed from the initial portfolio ω to portfolio ω', a distance d apart (where d is the number of assets changed in the portfolio). Let *p(d)* be the probability for any-given asset to be among the d assets that are changed by moving from ω to ω'. The autocorrelation coefficient, ρ*(d)*, for two portfolios a distance d apart is then given by:

35                              $$\rho(d) = 1-p(d). \tag{15}$$

The cost of an asset is unchanged if it is not one of the d assets that have been changed as the portfolio moved from ω to ω', and if it is not one of the K neighbors of any of the changed operations. These two events are statistically independent, and thus

$$\rho(d) = 1 - \left[1 - \frac{d}{N}\right]\left[1 - \frac{K}{N-1}\right]^d, \qquad (16)$$

from which it follows that

$$p(d) = \left[1 - \frac{d}{N}\right]\left[1 - \frac{K}{N-1}\right]^d. \qquad (17)$$

When $d = 1$ and there are no asset external dependencies $(K = 0)$, $\rho(d) = 1 - \frac{1}{N}$, which for N >> 1 is very close to 1; when every operation affects every other operation $(K = N)$, $\rho(d) = 0$.

When K increases, the landscape changes from "smooth" and single peaked to "rugged" and fully random. For low values of K the correlation spans the entire configuration space; the space is thus non-isotropic. As K increases, the configuration space breaks up into statistically equivalent regions, so the space as a whole becomes isotropic. See. Kauffman 1993, *supra*.

A related measure of landscape correlation, and one which can be used to compare landscapes, is the *correlation length*. The correlation length, l, of a landscape is defined by

$$l^{-1} = \sum_{d \geq 0} \rho(d). \qquad (18)$$

For a correlation coefficient which decays exponentially with distance, the correlation length is the distance over which the correlation falls to $1/K$ of its initial value. For the *NK* landscape

$$l = -\frac{1}{\ln\rho}. \qquad (19)$$

Consider an NK landscape with a moderately long correlation length and suppose that the search starts with a portfolio of average fitness 0.5 (for the rest of the discussion the fitness of the portfolio will be normalized to lie between 0 and 1). Then half of the 1-operation variant neighbors of the initial portfolio are expected to have a lower fitness, and half are expected to have higher fitness. More generally, half the of the portfolio variants at any distance d = 1, ... ,N away from the initial portfolio should be more fit and half should be

less fit. Since the landscape is correlated, however, nearby variants of the initial portfolio, those a distance 1 or 2 away, are constrained by the correlation structure of the landscape to be only slightly more or less fit than the starting configuration. In contrast, variants sampled at a distance well beyond the correlation length, l, of the landscape can have fitness very

5   much higher or lower than that of the initial portfolio.

Therefore, it is advantageous that, early in the search process from a poor or even average initial portfolio, the more fit variants are found most readily by searching far away on the landscape. But as the fitness increases, distant variants are found to be nearly average in the space of possible fitness - hence less fit - while nearby variants are likely to have

10   fitness similar to that of the current, highly fit, configuration. Thus, at this point distant search is less advantageous, while search is advantageously confined to the local region of the portfolio.

In another alternative, the landscape is represented using an annealed approximation, which is preferable for systems with disorder (i.e. randomly assigned properties) as is the

15   case with an NK model with K at least moderately large compared to N. The fitness ϕ^j are assigned by random sampling from $U(0, 1)$ (the uniform distribution). In evaluating the statistical properties of the NK landscape, first the entire landscape is sampled, and then some property on that landscape is measured. Repeated sampling and measuring on many landscapes then yields the desired aggregate statistics. To analytically approximate this

20   process of sampling and measuring the annealed approximation is preferably used. In an annealed approximation, the averaging over landscapes is done before measuring the desired statistic. Since the annealed approximation is sufficiently accurate for the purposes of determining optimum search parameters, it is the next alternative described.

As an example of an annealed approximation, assume a measurement of the average

25   of a product of four fitness along a connected walk is needed to determine optimal search parameters.. These fitness are labeled by θ_1, θ_2, θ_3, θ_4. If $P(\theta_1, \dots, \theta_S^N)$ is the probability distribution for an entire landscape (where S is the number of states, 2 in the case of a portfolio containing a particular asset or not), this average is calculated by the following.

$$\int \theta_1\theta_2\theta_3\theta_4 P(\theta_1, \cdots, \theta_{S^N} d\theta_1 \cdots d\theta_{S^N} = \int P(\theta_1, \theta_2, \theta_3, \theta_4 d\theta_1 d\theta_2 d\theta_3 d\theta_4. \tag{20}$$

30

This often difficult integral is, under the annealed approximation, is instead evaluated by the following.

$$\int P(\theta_1)\theta_1 P(\theta_2|\theta_1)\theta_2 P(\theta_3|\theta_2)\theta_3 P(\theta_4|\theta_3)\theta_4 d\theta_1 d\theta_2 d\theta_3 d\theta_4, \tag{21}$$

35

where $P(\theta|\theta')$ is the probability that a configuration has fitness θ conditioned on the fact that a neighboring configuration has fitness θ'.

According to the annealed approximation, the entire landscape is replaced by the joint probability distribution $P(\theta(\omega_i), \theta(\omega_j))$, where portfolios ω_i and ω_j are a distance one apart. For any particular landscape the probability that the fitness of a randomly chosen pair of portfolios a distance d apart have fitness θ and θ' the following.

$$P(\theta,\theta'|d) = \frac{\sum_{(\omega_i,\omega_j)_d} \delta(\theta-\theta(\omega_i))\delta(\theta'-\theta(\omega_j))}{\sum_{(\omega_i,\omega_j)_d} 1} \qquad (22)$$

where the notation $(\omega_i, \omega_j)_d$ requires that portfolios ω_i and ω_j are a distance d apart and δ is the Dirac delta function. The Dirac delta function is the continuous analog of the Kronecker delta function: $\delta(x)$ is zero unless $x = 0$ and is defined so that $\int_I dx \ \delta(x) = 1$ if the region of integration, I, includes zero. The full $P(\theta, \theta'|d)$ is advantageously we simplified and approximated by the following.

$$P(\theta(\omega_i),\theta(\omega_j)) \equiv P(\theta(\omega_i),\theta(\omega_j)|d = 1). \qquad (23)$$

For some landscape properties another approximation that can be used is the full $P(\theta(\omega_i), \theta(\omega_j)|d)$ distribution, as approximated by building up from $P(\theta(\omega_i), \theta(\omega_j))$. More accurate extensions of this annealed approximation may be obtained if $P(\theta(\omega_i), \theta(\omega_j)|d)$ is known.

From $P(\theta(\omega_i), \theta(\omega_j))$, both $P(\theta(\omega_i))$, the probability of a randomly chosen portfolio ω_i, having fitness $\theta(\omega_i)$, and $P(\theta(\omega_i)|\theta(\omega_j))$, the probability of a portfolio ω_i having fitness $\theta(\omega_i)$ given that a neighboring portfolio ω_j fitness $\theta(\omega_j)$ can be calculated. These probabilities are defined as

$$P(\theta(\omega_i)) = \int_{\infty}^{-\infty} P(\theta(\omega_i),\theta(\omega_j))d\theta(\omega_j), \qquad (24)$$

and

$$P(\theta(\omega_i)|\theta(\omega_j)) = \frac{P(\theta(\omega_i),\theta(\omega_j))}{P(\theta(\omega_j))}. \qquad (25)$$

Note that fitness ranges is assumed to be the entire real line, where fitness is not bounded from below, the ordering relationship amongst fitness is preserved, and extreme fitness are very unlikely.

For NK landscapes the following probability densities may be calculated exactly by the following known relationship.

$$P(\theta(\omega_i)) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\theta^2(\omega_i)}{2}\right],$$ (26)

$$P(\theta(\omega_i), \theta(\omega_j)) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{\theta^2(\omega_i) + \theta^2(\omega_j) - 2\rho\,\theta(\omega_i)\theta(\omega_j)}{2(1-\rho^2)}\right],$$ (27)

$$P(\theta(\omega_i)\,|\,\theta(\omega_j)) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[\frac{\theta(\omega_i) - \rho\,\theta(\omega_j))^2}{2(1-\rho^2)}\right];$$ (28)

where $\rho = 1 - K/N$ and where have assumed without loss of generality that the mean $\mu(\theta_i)$ and variance $\sigma^2(\theta_i)$ of the landscape are 0 and 1, respectively. This annealed approach approximates the NK technology landscape well when K/N ~1, that is, when $p \sim 0$, but can it is known that is can deviate when K/N ~ 0, *i.e.*, when $p \sim 1$. Equations (26) - (27) define a more general family of landscapes characterized by arbitrary ρ.

Since the effects of search at arbitrary distances d from a portfolio ω_i are required, $P(\theta(\omega_j)\,|\,\theta(\omega_i), d)$ is inferred from $P(\theta(\omega_i), \theta(\omega_j))$. This calculation is briefly described herein. To begin, note that $P(\theta(\omega_j)\,|\,\theta(\omega_i), d)$ is easily obtainable from $P(\theta(\omega_i)\,|\,\theta(\omega_j),\,|d)$ as

$$P(\theta(\omega_j)\,|\,\theta(\omega_i), d) = \frac{P(\theta(\omega_i),\theta(\omega_j)|d)}{P(\theta(\omega_i))}.$$ (29)

$P(\theta(\omega_i),\,|\,\theta(\omega_j)|d)$ is not known but it is related to $P(\theta(\omega_i), \theta(\omega_j)|s)$, the probability that a s-step random walk beginning at ω_i and ending at ω_j has fitness $\theta(\omega_i)$ and $\theta(\omega_j)$ at the endpoints of the walk. Each step of the random walk either increases or decreases the distance from the starting point by 1. $P(\theta(\omega_i), \theta(\omega_j)|s)$ is straightforward to calculate from

equation (27). $P(\theta(\omega_i), \theta(\omega_j)|d)$ is then obtained from $P(\theta(\omega_i), \theta(\omega_j)|s)$ by including the probability that a s-step random walk results in a net displacement of d-steps. The result of this calculation is that $P(\theta(\omega_j)|\theta(\omega_i), d)$ is Gaussian distributed with a mean and variance given by the following.

$$\mu(\omega_p, d) = \theta(\omega_i)\rho^d, \tag{30}$$

$$\sigma^2(\omega_p, d) = 1-\rho^{2d}. \tag{31}$$

In order to determine the relationship between search cost and optimal search distance on a landscape, the search problem is formulated preferably as dynamic programming problem. Each portfolio $\omega_i \in \Omega$ ($i = 1 \dots S^N$) is associated with a fitness θ_i. Portfolios at different locations in the landscape - and therefore at different distances from each other - have different Gaussian distributions corresponding to different $\mu(\omega_i, d)$ and $\sigma(\omega_i, d)$. A search cost, $c(d)$, is incurred every time a portfolio a distance d away from the current portfolio is sampled. The search cost $c(d)$ is a monotonic increasing function of d since more distant portfolios require greater changes to the current portfolio. For simplicity we take $c(d) = \alpha d$ (a linear relationship) but arbitrary functional forms for $c(d)$ are no more difficult to incorporate. The problem is to determine the optimal search distance at which to sample the landscape for improved portfolios. Note that since $E[\theta] \ll \infty$, by assumption, an optimal stopping rule exists for the search.

To determine the optimal distance at which to search for new portfolios, this alternative begins by denoting the current portfolio fitness by z. Supposing that one is considering sampling at a distance d. If $F_d(\theta)$ is the cumulative probability distribution of fitness at distance d, the expected fitness, $E(\theta|d)$, searching at distance d is given by

$$E(\theta|d) = -c(d) + \beta\left(z\int_{-\infty}^{z} dF_d(\theta) + \int_{z}^{\infty} \theta dF_d(\theta)\right). \tag{32}$$

where β is the discount factor. It can be the case that this expected-fitness discount factor is d-dependent since larger changes in the portfolio can require more effort but it is assumed, without loss of generality, that β is independent of d. The difference in fitness between searching at distance d and remaining with the current portfolio, $D_d(z)$, is given by:

$$D_d(z) = E(\theta|d)-z, \tag{33}$$

$$= -c(d) + \beta \left(z \int_{-\infty}^{z} dF_d(\theta) + \int_{z}^{\infty} \theta dF_d(\theta) \right) - z \qquad (34)$$

$$= -c(d) - (1-\beta)z + \beta \int_{z}^{\infty} (\theta - z) dF_d(\theta). \qquad (35)$$

$D_d(z)$ is a monotonic decreasing function of z which crosses zero at $z_c(d)$, determined by $D_d(z_c(d)) = 0$. For $z < z_c(d)$ it is preferable to sample a new portfolio ω_j since $D_d(z)$ is positive. If $z > z_c(d)$ it is preferable to remain with the current portfolio ω_i because $D_d(z)$ will be negative and the cost will outweigh the potential gain. The zero-crossing value $z_c(d)$ thus plays the role of a known reservation price. The reservation price at distance d is determined from the integral equation:

$$c(d) + (1-\beta)z_c(d) = \beta \int_{z_c(d)}^{\infty} (\theta - z_c(d)) dF_d(\theta). \qquad (36)$$

From the above equation, it can be seen that, as expected, the reservation price decreases with greater search cost.

The optimal search strategy on the landscape can be characterized by Pandora's Rule: if a portfolio at some distance is to be sampled, it preferably is a portfolio at the distance with the highest reservation price. The search preferably terminates and remains at the current portfolio whenever the current fitness is greater than the reservation price of all distances.

In the case where fitness at distance d are Gaussian distributed, the above equation can be formulated as. (For clarity the d dependence of z_c has been omitted)

$$c(d) + (1-\beta)z_c = \frac{\beta}{\sqrt{2\pi}} \int_{z_c}^{\infty} \frac{d\theta}{\sigma(\omega_i, d)} (\theta - z_c) \exp\left[-\frac{(\theta - \mu(\omega_i, d))^2}{2\sigma^2(\omega_i, d)} \right], \qquad (37)$$

$$= \frac{\beta}{\sqrt{2\pi}} \int_{0}^{\infty} \frac{du}{\sigma(\omega_1, d)} u \exp\left[-\frac{(u + z_c - \mu(\omega_1, d))^2}{2\sigma^2(\omega_1, d)} \right]. \qquad (38)$$

From the indefinite integral

$$\int \frac{du}{b} u \exp\left[-\frac{(u-a)^2}{2b^2}\right] = -b\exp\left[-\frac{(u-a)^2}{2b^2}\right] + a\sqrt{\frac{\pi}{2}}erf\left[-\frac{u-a}{\sqrt{2}b}\right], \qquad (39)$$

the following is found.

$$\int_0^\infty \frac{du}{b} u \exp\left[-\frac{(u-a)^2}{2b^2}\right] = b\exp\left[-\frac{a^2}{2b^2}\right] + a\sqrt{\frac{\pi}{2}}erfc\left[-\frac{a}{\sqrt{2}b}\right] \qquad (40)$$

where $erf[\cdot]$ is the error function and $erfc[\cdot] = 1 - erf[\cdot]$ is the complimentary error function. The error function $erf(x)$ is defined as $\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2} dt$ and the complimentary error function, $erfc(x)$ is defined as $\frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2} dt$. From these definitions it is easy to show that $erf(x) + erfc(x) = 1$, $erf(\infty) = 1$ and $erfc(-x) = 2 - erfc(x)$. With this result the equation determining the reservation price now reads:

$$c(d) + (1-\beta)z_c = \beta\left(\frac{\mu(\omega_1,d)-z_c}{2}erfc\left[-\frac{\mu(\omega_1,d)-z_c}{\sqrt{2}\sigma(\omega_1,d)}\right] + \frac{\sigma(\omega_1,d)}{\sqrt{2\pi}}\exp\left[-\frac{(\mu(\omega_1,d)-z_c)^2}{2\sigma^2(\omega_1,d)}\right]\right). \qquad (41)$$

To simplify the appearance of this equation, it is written using the dimensionless variable

$$\delta = \frac{z_c - \mu(\omega_1,d)}{\sqrt{2}\sigma(\omega_1,d)}, \qquad (42)$$

in terms of which $z_c = \sqrt{2}\sigma(\omega_1,d)\delta + \mu(\omega_1,d)$. The dimensionless reservation price δ is then determined by

$$\frac{\sqrt{2}(c(d) + (1-\beta)\mu(\omega_1,d))}{\sigma(\omega_1,d)} = \beta\left(\frac{\exp[-\delta^2]}{\sqrt{\pi}} - \delta erfc[\delta]\right) - 2(1-\beta)\delta, \qquad (43)$$

$$A(\omega_1, d) \equiv \frac{\sqrt{2}(c(d) + (1-\beta)\mu(\omega_1, d)}{\sigma(\omega_1, d)}, \tag{44}$$

Defining

$$= \beta\left(\frac{\exp[-\delta^2]}{\sqrt{\pi}} + \delta \, erfc[-\delta]\right) - 2\delta. \tag{45}$$

the equation which is solved for δ is therefore:

$$A(\omega_1, d) \equiv \beta\left(\frac{\exp[-\delta^2]}{\sqrt{\pi}} + \delta \, erfc[-\delta]\right) - 2\delta. \tag{46}$$

The explicit ω_1 and d dependence of A is obtained by substituting equations (30) and (31) into equation (45). Equation (46) is the central equation determining the reservation price $z_c(d)$.

The optimal search distance, d^*, is now determined as

$$d^* \equiv arg\,max_d z_c(d). \tag{47}$$

where the d-dependence of $z_c(d)$ is implicitly determined by Equation (46). As a function of d, z_c is smoothly behaved with a single maximum so that d^* is the integer nearest to the d which solves $\partial_d z_c = 0$. Next, the equation which d^* satisfies is found.

To begin, recall the definition of δ given in Equation (42). Taking the d derivative of δ yields

$$\partial_d z_c = \sqrt{2}\left((\delta\partial_d\sigma(\omega_1, d) + \sigma(\omega_1, d)\partial_d\delta\right) + \partial_d\mu(\omega_1, d). \tag{48}$$

The partial derivatives $\partial_d\mu$ and $\partial_d\sigma$ are given by

$$\partial_d\mu(\omega_1, d) = d\theta(\omega_1)\rho^{d-1}, \tag{49}$$

$$\partial_d \sigma(\omega_1, d) = -2d\rho^{2d-1}, \tag{50}$$

respectively, and $\partial_d \delta$ is next expressed in terms of these known quantities. Differentiating equation (34) with respect to d yields.

$$\partial_d \sigma = \frac{\partial_d A(\omega_1, d)}{\beta \, erfc[-\delta] - 2}. \tag{51}$$

(assuming β is not d-dependent). Thus d^* is determined by

$$0 = \sqrt{2} \left(\delta \partial_d \sigma + \frac{\sigma \partial_d A}{\beta \, erfc[-\delta] - 2} \right) + \partial_d \mu. \tag{52}$$

Using the definition of A in equation (45) its derivative is easily found as

$$\partial_d A = \frac{\sqrt{2}}{\sigma} \partial_d c + \frac{\sqrt{2}(1-\beta)}{\sigma} \partial_d \mu - \frac{A}{\sigma} \partial_d \sigma. \tag{53}$$

Substituting this result the following is found.

$$0 = \sqrt{2} \left(\delta \partial_d \sigma + \frac{\sqrt{2}\partial_d c + \sqrt{2}(1-\beta)\partial_d \mu - A\partial_d \sigma}{\beta \, erfc[-\sigma] - 2} \right) + \partial_d \mu, \tag{54}$$

which can be rearranged to give

$$0 = 2\partial_d c + \sqrt{2}(\beta \delta \, erfc[-\delta] - 2\delta - A)\partial_d \sigma + \beta(erfc[-\sigma] - 2)\partial_d \mu. \tag{55}$$

Finally, we use equation (46) to simplify this to,

$$\frac{2}{\beta} \partial_d c = \sqrt{\frac{2}{\pi}} \exp[-\delta^2]\partial_d \sigma + erfc[\delta] \partial_d \mu \tag{56}$$

where $\partial_d \mu$ and $\partial_d \sigma$ are given in equations (49) and (50).

Once this distance, d^*, is known, in step 430 the method searches for optimal portfolios by making steps to neighboring portfolios at the optimal searching distance. Further, other parameters of the fitness landscape search can be optimized as described in

the above reference. For example, the method illustrated in FIG. 4 can also be used to determine indifference surfaces as part of the method illustrated in FIG. 3.

In further alternative embodiments, method 100 can repeat the data gathering step 104 in view of the landscape complexity (or other landscape parameter) determined in the generation of clusters and indifference surfaces. Data gathering can be repeated in an optimized manner to determine the important parameters of preference landscapes with increased accuracy according to their observed previously approximate ruggedness. Additional questions could be asked of the customer to choose among characteristics which are more closely aligned to portfolios of goods which may have, for example, a greater preference or a lower VAR. After these additional preferences are solicited, the rest of the process can be repeated to repartition the customers, create new indifference surfaces, and optimize the portfolio synthesized.

Finally, the synthesized portfolios are offered to the customers. Although this step is outside the scope of this invention, being carried out by profit-maximizing economic actors, such actors have increased assurance that the portfolios synthesized according to the present invention will be both profitable to offer and satisfactory to customers.

The present invention also includes systems for gathering customer data and providing optimized portfolios. Fig. 5 illustrates exemplarily such system 500 in conjunction with which the embodiments of the present invention can be implemented. User devices 502, *inter alia,* gather preference data from the customers and return optimized portfolio offerings. These user devices include, but are not limited to, computer terminals, handheld personal data assistants, personal computers, telephones. Alternatively, user devices can be directly attached to server systems 504.

Server systems 504 perform the methods of partitioning the customers into a plurality of clusters according to the preference data, generating indifference surfaces based on their preferences, and synthesizing a portfolio for them. The server computers include CPUs, dynamic memory accessible by the CPU for retrieving instructions and data, permanent storage, such as tape devices, disc drives and CD-ROMs readers, and network interfaces for communicating to user devices. When computer instructions implementing the methods of the present invention are loaded into the directly accessible dynamic memory of the server systems, their CPUs are commanded to perform the methods of this invention. The server systems include storage devices that can be loaded with historical data pertaining to goods, services or financial instruments.

Source programs implementing the above-described methods of this invention can be written in convenient computer languages by artisans of average skill in view of the previous descriptions. Computer instructions generated by such source programs can be

stored on computer readable media for loading into server computer storage, or can be transmitted over a network to such storage.

Communication network 506 serves to communicate preference data from the user device 502 to server systems 504. The communications network includes, but is not limited

5   to a packet switched data networks, a local or wide area network, or the Internet.

Also attached to the communications network are business systems 508. The server computers can cooperate with the business systems to provide related to the feasibility of candidate portfolios and to arrange for providing optimum portfolios. The feasibility data can include technologic, economic, or historic data as described above. For example, in the

10 case of insurance services, the business systems are of insurers which contain necessary historical risk-of-loss data and policy information. The server and insurer business systems can cooperate to make determined optimum portfolios of insurance services available to users at the user devices. In the case of financial instruments, the business systems can include exchange and brokerage systems. In the case of goods, the business systems can be

15 those of the manufacturers of the goods.

While the above invention has been described with reference to certain preferred embodiments, the scope of the present invention is not limited to these embodiments. One skilled in the art may find variations of these preferred embodiments which, nevertheless, fall within the spirit of the present invention, whose scope is defined by the claims set forth

20 below. All references cited herein are incorporated herein by reference in their entirety and for all purposes to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes.

25

30

35

What is claimed is:

1.   A method for dynamically synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, said method
5   comprising the steps of:
        gathering preference data from a plurality of customers, wherein the preference data is responsive to the preference of each individual for individual goods, services or financial instruments,
        partitioning the customers into a plurality of clusters of customers according to said
10   preference data, wherein the customers of each cluster have similar preferences, and
        synthesizing at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments included in each portfolio are based on the preferences of the cluster of customers. .

15   2.   The method of claim 1 wherein said step of gathering preference data further comprises querying the customers.

        3.   The method of claim 1 wherein said step of gathering preference data further comprises searching databases containing data related to customer behavior.
20

        4.   The method of claim 1 wherein said step of partitioning the customers further comprises performing a k-means clustering method or a multidimensional scaling method.

25   5.   The method of claim 4 wherein said step of partitioning the customers further comprises performing an adaptive dissimilarity partitioning method.

        6.   The method of claim 5 wherein said adaptive dissimilarity partitioning method selects a measure of dissimilarity from a chosen family of dissimilarity measures,
30   wherein the clusters of customers are defined by the selected measure of dissimilarity.

        7.   The method of claim 5 wherein said adaptive dissimilarity partitioning method further comprises performing a genetic algorithm.

35

8.   The method of claim 1 wherein said step of synthesizing further comprises generating at least one utility surface for each cluster of customers, wherein a point on the utility surface indicates the utility of the portfolio represented by the point for the cluster of customers, and wherein the utility surface is based on the preferences of the cluster of

5 customers.

9.   The method of claim 1 wherein said step of synthesizing further comprises generating at least one indifference surface for each cluster of customers, wherein a point on the indifference surface indicates the preference of the cluster of customers for the portfolio

10 represented by the point, and wherein the indifference surface is based on the preference of the cluster of customers.

10.   The method of claim 9 wherein the indifference surface is modeled by a parameterized model of fitness landscapes.

15

11.   The method of claim 10 wherein the parameterized model is an NK model or an order-P spin-glass model.

12.   The method of claim 9 wherein the indifference surface has a plurality of

20 peaks, and wherein the indifference surface is characterized by one or more parameters, including a number of peaks on the indifference surface, or an expected number of steps to a peak from any point on the indifference surface, or a rate of decrease in a number of directions of increase as a peak of the indifference surface is approached, or a number of different peaks that can be approached from a single point on the indifference surface by

25 adaptive walks proceeding only in directions of increase, or a correlation structure of the indifference surface.

13.   The method of claim 12 wherein the correlation structure of the indifference surface is measured by a correlation between two points on the indifference surface as a

30 function of a distance between of two points.

14.   The method of claim 9 wherein said step of synthesizing further comprises searching the indifference surface for portfolios an optimal distance away for indication of relatively greater preference.

35

15. The method of claim 12 further comprising gathering further preference data from the customers in order to determine the parameters of the parameterized model of the indifference surface.

5   16. The method of claim 1, wherein the individual services comprise insurance services.

17. The method of claim 1, wherein the financial instruments comprise stocks or bonds.

10

18. The method of claim 1 further comprising offering at least one synthesized portfolio to at least one customer of a corresponding cluster.

19. A method for dynamically synthesizing portfolios comprising a plurality of
15 individual goods, individual services or individual financial instruments, said method comprising the steps of:

        gathering preference data from a plurality of customers, wherein the preference data is responsive to a preference of each customer for individual goods, services or financial instruments,

20   partitioning the customers into a plurality of clusters of customers according to said preference data, wherein the customers of each cluster have similar preferences,

        generating at least one indifference surface for each cluster of customers, wherein a point on the indifference surface indicates the preferences of the customers of the cluster for the portfolio represented by the point, and wherein the indifference surface is based on the
25 preferences of the customers in the cluster, and

        synthesizing at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments included in each synthesized portfolio are based on the preference of the cluster of customers, and wherein said synthesizing further comprises searching the indifference surface for portfolios
30 indicating a relatively greater preference.

20. The method of claim 19 wherein said searching further comprises repetitively proceeding by steps from a current portfolio to a next portfolio, wherein the next portfolio becomes a succeeding current portfolio if the next portfolio has a relatively
35 increased preference compared to the current portfolio.

21.   The method of claim 20 wherein said searching further comprises selecting a size of the step responsive to ruggedness of the indifference surface.

22.   The method of claim 21 wherein said selecting further comprises selecting a smaller size for a relatively rugged indifference surface compared to the size selected for a relatively smooth indifference surface.

23.   The method of claim 19 said searching further comprises performing a genetic-algorithm search method.

24.   The method of claim 23 wherein the genetic-algorithm search method evolves a population of portfolios on the indifference surface.

25.   The method of claim 23 wherein parameters of the genetic-algorithm search method are selected to be responsive to the ruggedness of the indifference surface.

26. The method of claim 19 further comprises providing at least one fitness surface, and wherein said searching further comprises searching the indifference surface and the further fitness surface simultaneously for optima.

27.   The method of claim 26 wherein said searching further comprises searching for at least one Pareto optimum dominating optima of the further fitness surface and optima of the indifference surface.

28.   The method of claim 26 wherein the portfolio comprises goods, and wherein the further fitness surface is responsive to an economic or to a technological feasibility of a candidate portfolio.

29.   The method of claim 26 wherein the portfolio comprises financial instruments, and wherein the further fitness surface is responsive to a feasibility of acquisition or of divestiture of the financial instruments.

30.   The method of claim 26 wherein the portfolio comprises financial instruments, and wherein the further fitness surface is responsive to value at risk of portfolios of financial instruments.

31.   The method of claim 30 further comprising determining a value at risk of the financial instruments from historical data.

32.   The method of claim 26 wherein the portfolio comprises insurance services, and wherein the further fitness surface is responsive to a risk of loss of goods insured by the insurance services.

33.   The method of claim 32 further comprising determining the risks of loss from historical data.

34.   The method of claim 26 wherein the portfolio comprises insurance services, and wherein the further fitness surface is responsive to value at risk of goods insured by the insurance services.

35.   The method of claim 19 wherein said step of gathering preference data from a plurality of customers further comprises
        querying customers by transmitting query messages to customers, and
        obtaining said preference data from response messages received from customers in response to transmitted messages.

36.   A system for dynamically synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, said system comprising:
        at least one user device for gathering preference data from a plurality of customers;
        at least one server computer configured by computer instructions to cause the server computer to
                gather preference data from a plurality of customers, wherein the preference data is responsive to the preference of each individual for individual goods, services or financial instruments, and to
                partition the customers into a plurality of clusters of customers according to said preference data, wherein the customers of each cluster have similar preferences, and to
                synthesize at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments included in each synthesized portfolio are based on the preferences of the customers of the cluster, and

at least one communications network for communicating between the user devices and the server computers.

38. The system of claim 36 wherein the server computer contains instructions
5 which further cause the server computer to, as part of said step of synthesizing, generate at least one indifference surface for each cluster of customers, wherein a point on the indifference surface indicates a preference of the customers of the cluster for the portfolio represented by the point, and wherein the indifference surface is based on the preference of the customers in the cluster.

10

39. The system of claim 36 wherein the server computer instructions further cause the server computer to offer at a user device at least one synthesized portfolio to at least one customer of the corresponding cluster.

15   40, The system of claim 36 further comprising at least one business system for providing feasibility data for portfolios, and wherein the communication network communicates between the server computer and the business systems.

41. The system of claim 40 wherein the business systems are means for
20 synthesizing portfolios.

42. The system of claim 36 wherein the communications network is a packet switched data network.

25   43. The system of claim 36 wherein the communications network comprises the Internet or an intranet..

44. A server system for dynamically synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, said
30 server system comprising:
    at least one CPU,
    memory dynamically accessible by the CPU, wherein the memory is configured with computer instructions for causing the CPU to
        gather preference data from a plurality of customers at user devices, wherein
35 the preference data is responsive to the preference of each individual for individual goods, services or financial instruments, and to

partition the customers into a plurality of clusters of customers according to said preference data, wherein the customers of each cluster have similar preferences, and to synthesize at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments

5   included in each synthesized portfolio are based on the preferences of the customers of the cluster.

45.   The server system of claim 44 wherein the computer instructions further cause the CPU to, as part of said step of synthesizing, generate at least one indifference

10   surface for each cluster of customers, wherein a point on the indifference surface indicates the preferences of the customers of the cluster for the portfolio represented by the point, and wherein the indifference surface is based on the preferences of the customers in the cluster.

46.   The server system of claim 44 wherein the computer instructions further

15   cause the CPU to offer at a user device at least one synthesized portfolio to at least one customer of the corresponding cluster.

47.   The server system of claim 44 wherein the computer instructions further cause the CPU to communicate with a business data system in order to obtain portfolio

20   feasibility data or to provide synthesized portfolios to customers.

48.   The server system of claim 44 further comprising an interface for communication with a network.

25   49.   A system for dynamically synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, said system comprising:

means for gathering preference data from a plurality of customers at user devices, wherein the preference data is responsive to the preference of each individual for individual

30   goods, services or financial instruments,

means for partitioning the customers into a plurality of clusters of customers according to said preference data, wherein the customers of each cluster have similar preferences, and

means for synthesizing at least one portfolio for each of the clusters of customers,

35   wherein the individual goods, individual services, or individual financial instruments

included in each synthesized portfolio are based on the preferences of the customers of the cluster.

50.   The system of claim 49 wherein the means for synthesizing further

5   comprises means for generating at least one indifference surface for each cluster of customers, wherein a point on the indifference surface indicates the preferences of the customers of the cluster for the portfolio represented by the point, and wherein the indifference surface is based on the preferences of the customers in the cluster.

10   51.   The system of claim 49 further comprising means for offering at a user device at least one synthesized portfolio to at least one customer of the corresponding cluster.

52.   A computer readable medium comprising encoded computer instructions for

15   causing a computer to perform a method for dynamically synthesizing portfolios comprising a plurality of individual goods, individual services or individual financial instruments, said method comprising:

gathering preference data from a plurality of customers, wherein the preference data is responsive to the preference of each individual for individual goods, services or financial

20   instruments,

partitioning the customers into a plurality of clusters of customers according to said preference data, wherein the customers of each cluster have similar preferences, and

synthesizing at least one portfolio for each of the clusters of customers, wherein the individual goods, individual services, or individual financial instruments included in each

25   synthesized portfolio are based on the preferences of the customers of the cluster.

30

35

100

```
┌─────────────────┐
│ Gather customer │ ─── 102
│      data       │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│   Partition     │ ─── 104
│ customers into  │
│    clusters     │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│    Generate     │
│   indifference  │ ─── 105
│ surfaces for each│
│     cluster     │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│   Synthesize    │ ─── 106
│  portfolios for │
│   each cluster  │
└─────────────────┘
```

FIG. 1

FIG. 2

300

FIG. 3

400

```
┌─────────────────┐
│ Enter determined │
│ fitness landscape│  ～ 410
│ of objective ot be│
│   optimized     │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│   Determine     │
│    optimal      │  ～ 420
│   searching     │
│    distance     │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│   Search for    │
│ optimal portfolios│  ～ 430
│  at the optimal │
│   searching     │
│    distance     │
└─────────────────┘
```

FIG. 4

560

User devices

← 502 ~

Communications
network

508 Business
systems

504

506

Server systems
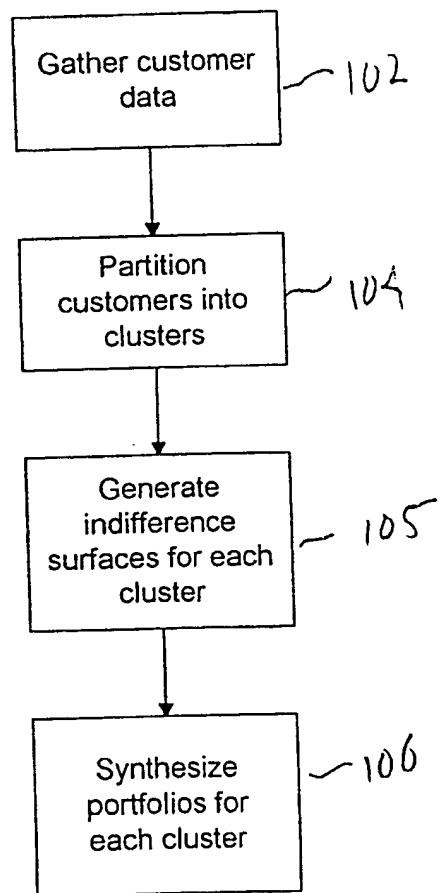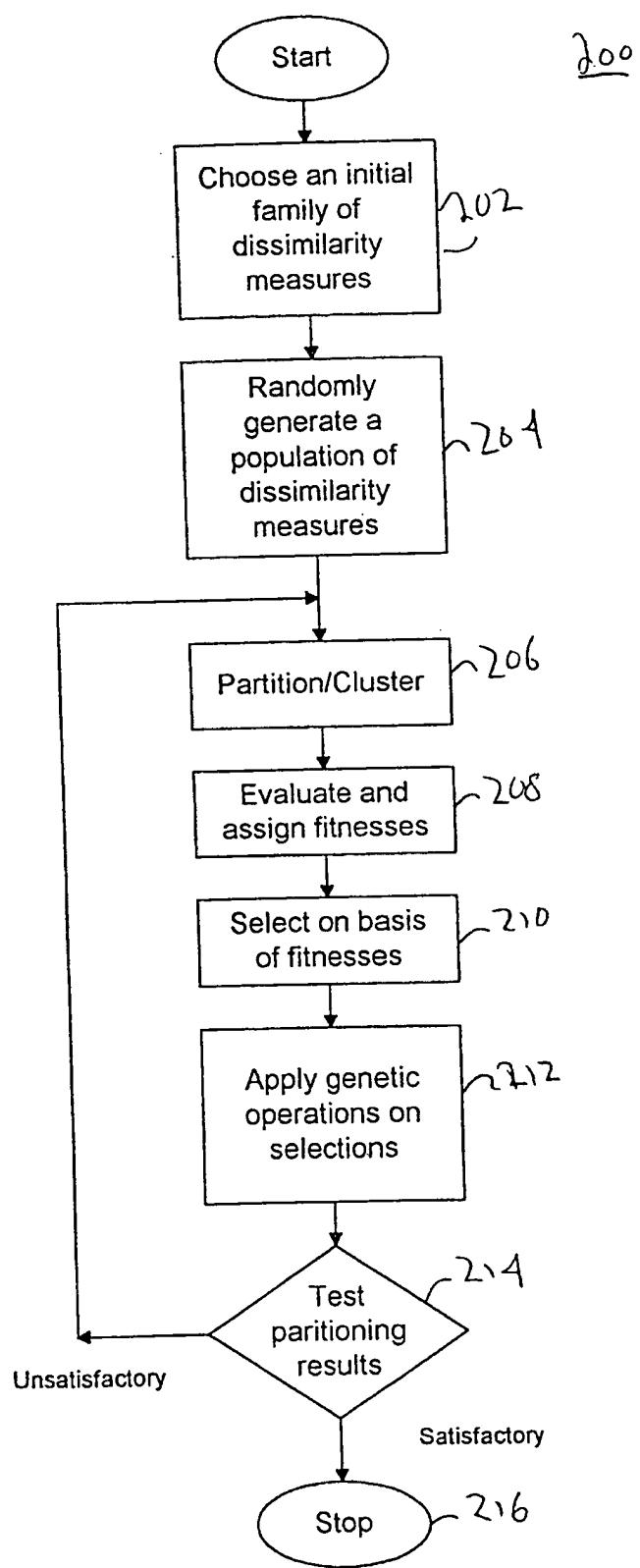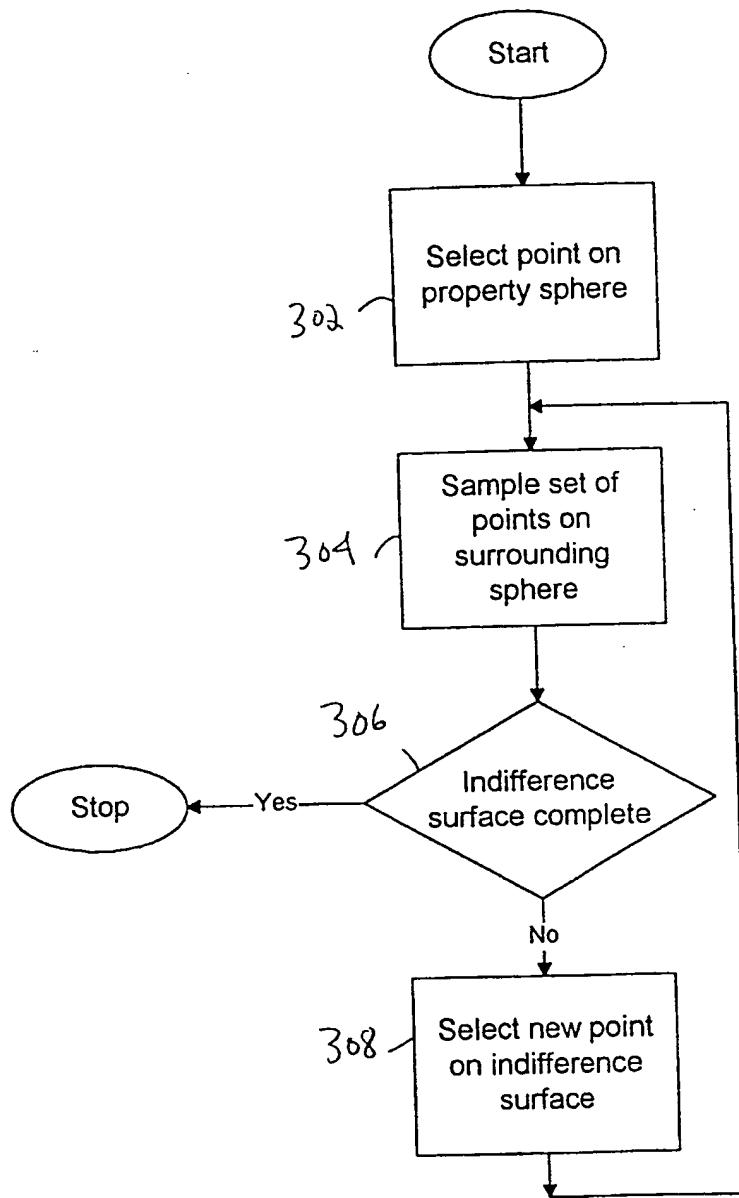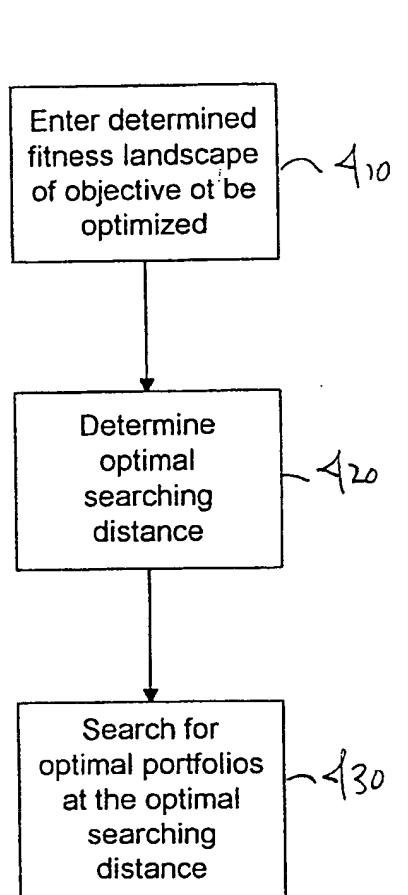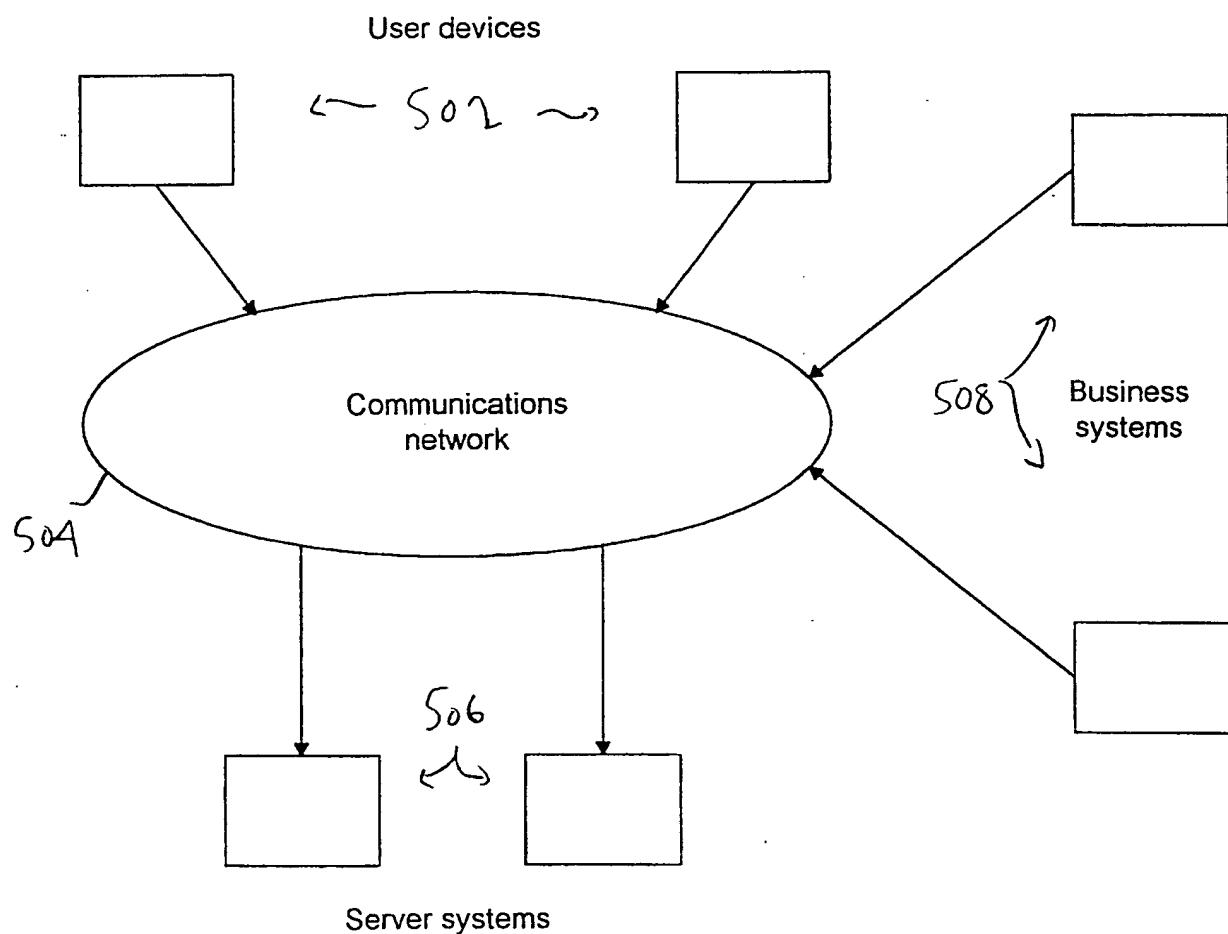
FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/18632

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7)   :G06F 17/60
US CL   : 705/36

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :   705/36

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, INTERNET, DIALOG
PORTFOLIO, MULTIDIMENSIONAL SCALING, PREFERENCE, GENEREC ALGORITHM

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BUILDING MARKET STRUCTURES FROM CUSTOMER PREFERENCES; COOPER LEE G.; INOUE, AKIHIRO; JOURNAL OF MARKETING RESEARCH,; AUGUST 1996; V33, N3, P293 (14)	1-52
Y	A FEATURE-BASED APPROACH TO MARKET SEGMENTATION VIA OVERLAPPING K-CENTROIDS CLUSTERING. CHATURVEDI, ANIL ET AL.; AUGUST 1997; JOURNAL OF MARKETING RESEARCH, V34, N3, P370(8)	1-52
Y	INTEGRATING BUSINESS STRAGEGY AND CAPITAL ALLOCATION: AN APPLICATION OF MULTI-OBJECTIVE DECISION MAKING; WALLS, MICHAEL; ENGINEERING ECONOMIST; SPRING 1995; V40 N3 PP: 247- 266.	1-52

[X] Further documents are listed in the continuation of Box C.   [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 SEPTEMBER 2000	**0 8 DEC 2000**

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks	
Box PCT	JAMES TRAMMELL
Washington, D.C. 20231	
Facsimile No. (703) 305-3230	Telephone No. (703) 305-9768

Form PCT/ISA/210 (second sheet) (July 1998)*

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	THE JAPANESE CORPORATE NETWORK: A BLOCKMODEL ANALYSIS; GERLACH, MICHAEL L. AMINISTRATIVE SCIENCE QUARTERLY; MARCH 1992; V37N1 PP: 105-139.	1-52
Y	A CLUSTERWISE REGRESSION METHOD FOR SIMULTANEAUS FUZZY MARKET STRUCTURING AND BENEFIT SEGMENTATION; WEDEL, MICHEL; JOURNAL OF MARKETING RESEARCH, NOVEMBER 1991;V28, N4, P385 (12)	1-52
Y	SOLVING MARKETING OPTIMZATION PROBLEMS USING GENETIC ALOGRITHMS; HURLEY, S. ET AL.; EUROPEAN JOURNAL OF MARKETING; 1995; V29N4, PP: 39-56.	7, 23-25